ROUGH-CUT CAPACITY PLANNING IN MULTIMODAL FREIGHT TRANSPORTATION NETWORKS

DISSERTATION

Robert B. Hartlage, Major, USAF

AFIT/DS/ENS/12-03

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/DS/ENS/12-03

Rough-Cut Capacity Planning in Multimodal Freight Transportation Networks

DISSERTATION

Presented to the Faculty

Department of Operational Sciences

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

Robert B. Hartlage, B.S., M.S.Eng., M.S.

Major, USAF

September 2012

AFIT/DS/ENS/12-03

Rough-Cut Capacity Planning in Multimodal Freight Transportation Networks

Robert B. Hartlage, B.S., M.S.Eng., M.S.
Major, USAF

Approved:

Jeffery D. Weir — 30 Aug 2012

Jeffery D. Weir, PhD (Chairman) — Date

Raymond R. Hill — 20 Aug 2012

Raymond R. Hill, PhD (Member) — Date

Kenneth M. Hopkinson — 20 Aug 2012

Kenneth M. Hopkinson, PhD (Member) — Date

Accepted:

MU Thomas — 6 Sep 2012

M. U. Thomas, PhD
Dean, Graduate School of
Engineering and Management — Date

To Granddad, who taught me many important life lessons through colorful metaphor…and no Granddad, they're not grinding me down.

AFIT/DS/ENS/12-03

Abstract

A main challenge in transporting cargo for United States Transportation Command (USTRANSCOM) is in mode selection or integration. Demand for cargo is time sensitive and must be fulfilled by an established due date. Since these due dates are often inflexible, commercial carriers are used at an enormous expense, in order to fill the gap in organic transportation asset capacity. This dissertation develops a new methodology for transportation capacity assignment to routes based on the Resource Constrained Shortest Path Problem (RCSP). Routes can be single or multimodal depending on the characteristics of the network, delivery timeline, modal capacities, and costs. The difficulty of the RCSP requires use of metaheuristics to produce solutions. An Ant Colony System to solve the RCSP is developed in this dissertation. Finally, a method for generating near Pareto optimal solutions with respect to the objectives of cost and time is developed.

Acknowledgements

I first acknowledge my Lord and Savior Jesus Christ for this accomplishment. When I have enjoyed success it is to His glory and when I have failed it is a testament only to my own shortcomings.

Completion of this dissertation would have been impossible without my wife's support and willingness to take the family reins while I've been distracted for the past three years. At the end of every day I'm "daddy" and "hubby", titles I will always wear with greater pride than any other achievement.

I'd like to thank my parents for demonstrating to me what it means to succeed in marriage and in raising a family. My Dad is the wisest man I'll ever know. His sacrifice for his family have taught me lessons I will always try to emulate. My Mom invested the years of her youth in her family, selflessly keeping it all together behind the scenes. As Mom's often times do, she always believed I could.

I am deeply indebted to my dissertation advisor, Dr. Jeff Weir whose "even keel" kept me from sinking too low or rising too high. His guidance throughout this arduous process have been invaluable. Someday I hope to be even half as good of an professor, advisor, and researcher as Dr. Weir.

Thanks to Dr. Hill without whom I'd never have pursued this PhD and to Dr. Hopkinson for insightful questions that have resulted in several research improvements.

Col Andrew and the United States Air Force Academy Math Department deserve my sincerest thanks for the opportunity to pursue this PhD.

Finally, to PhD-12 and the "Unofficial AFIT Homebrew Club", thanks for the camaraderie and beer-snobbery...alas fellas, it's time for us to go get real jobs!

Robert B. Hartlage

Table of Contents

Page

Abstract vi

Acknowledgements vii

List of Figures xi

List of Tables xii

I. Introduction 1
- 1.1 Background & Motivation 1
- 1.2 Problem Description, Themes, & Gaps 2
  - 1.2.1 Themes 3
  - 1.2.2 Gaps 4
- 1.3 Problem Statement 5
  - 1.3.1 Research Questions 6
- 1.4 Research Objectives 7
  - 1.4.1 Research Objectives & Contributions 7
- 1.5 Dissertation Overview 8

II. Literature Review 9
- 2.1 Literature Review Strategy 9
- 2.2 Overview of the Literature 9
- 2.3 Multimodal Freight Transportation 13
  - 2.3.1 Intermodal Transportation 14
- 2.4 Multi-Layered Networks 15
- 2.5 Hierarchical Production Planning 15
  - 2.5.1 Aggregate Production Planning 15
  - 2.5.2 Rough-Cut Capacity Planning 16
  - 2.5.3 Master Production Scheduling 17
- 2.6 Capacity Planning in Multimodal Freight Transportation 17
- 2.7 Shortest Path and Resource Constrained Shortest Path Problems 18

Page

III. Rough-Cut Capacity Planning using Resource-Constrained Shortest Paths 19
 3.1 Introduction 19
  3.1.1 Assumptions 27
 3.2 Modeling Approach 27
  3.2.1 Single Mode Formulation 30
  3.2.2 Multimodal Formulation 32
 3.3 Example Problem Formulation 39
  3.3.1 An Example Formulation 39
  3.3.2 Mode Selection & Integration 47
 3.4 Results 48
  3.4.1 Parameter Selection & MPS File Generation Issues 49
  3.4.2 Table of Results for Successfully Generated Problems 49
 3.5 Contributions and Future Work 50

IV. An Ant Colony System for Solving the Resource Constrained Shortest Path Problem 52
 4.1 Introduction 52
  4.1.1 Solving the RCSP 53
  4.1.2 Constructive and Local Search Metaheuristics 53
  4.1.3 Swarm Intelligence and Ant Colony Optimization 54
  4.1.4 The Basic Ant Colony Optimization Metaheuristic 55
 4.2 Modeling Approach 58
  4.2.1 Selecting an Ant Colony Metaheuristic 58
  4.2.2 Extension of ACS to Solve RCSP Problem . . 61
 4.3 Testing the ACS Metaheuristic 63
  4.3.1 Selection of Parameter Values 63
 4.4 Results 69
  4.4.1 Small Example Problems 69
  4.4.2 Larger Example Problems 70
 4.5 Conclusions and Future Work 71

V. Finding Near Efficient Solutions to Lexicographic, Multi-Objective, Resource Constrained Shortest Path Problems 73
 5.1 Introduction 73
  5.1.1 Notions of Optimality 74
  5.1.2 Problem Statement 75
  5.1.3 Review of the Literature 75
 5.2 Methodology 83

Page

5.2.1 Finding Near-Efficient Solutions 84
5.2.2 Distance between Solutions and Coverage of the Pareto Front 85
5.2.3 Generalized Range Variance for Lexicographic Objective Ordering 86
5.2.4 The Value of Dominated Solutions 87
5.2.5 Methodology Development 87
5.3 Example Problems 92
5.3.1 Randomly Generated Examples 92
5.4 Conclusions and Future Work 98
5.4.1 Conclusions 98

VI. Summary & Conclusions 100

VII. Appendix 1: FCD Experiment Results For ACS Testing 102

Bibliography 106

List of Figures

Figure		Page
2.1.	Overview of Reviewed Literature Sources	10
2.2.	Overview of Reviewed Literature Sources	11
3.1.	An Example Network: G^1	28
3.2.	Reweighted Network	29
3.3.	An Example Network: G^1	40
3.4.	Graph for G^2	40
3.5.	Graph for G^3	40
3.6.	Intermodal Graphs of H^1, H^2, H^3, and H^4	41
3.7.	Multimodal Graph	42
4.1.	Percent over optimal for actual vs. predicted	66
4.2.	Run Time for actual vs. predicted	68

List of Tables

Table		Page
2.1.	Features of this dissertation versus literature's research	12
2.2.	Overview of this dissertation	13
3.1.	Problem Generation and Solution Times	50
4.1.	Parameter Value Ranges	64
4.2.	Fit Statisitcs for % over optimal	66
4.3.	ANOVA Table for % over optimal	66
4.4.	ACS Performance during Initial Experimentation	67
4.5.	Fit Statistics for Run Time	67
4.6.	ANOVA Table for Run Time	68
4.7.	Final ACS Parameters Maximizing Desirability	68
4.8.	Max Desirability vs. Parameter Select Expt Settings	70
4.9.	Mean and $\hat{\sigma}$ for Large Experiments	71
5.1.	Table of UTE Factors	74
5.2.	Small Experiment Parameters	96
5.3.	Multiobjective Lexicographic Order & Time Solution Results	97
5.4.	Multiobjective Cost & Time Solution Results	97
7.1.	FCD Experimental Results Part 1	103
7.2.	FCD Experimental Results Part 2	104
7.3.	FCD Experimental Results Part 3	105

Rough-Cut Capacity Planning in Multimodal Freight Transportation Networks

I. Introduction

1.1 Background & Motivation

Transportation networks are prevalent in nearly every aspect of our society and as of the most recent economic census, transportation revenues in the U.S. represent a nearly $640 billion segment of the economy [1]. Domestic and international commerce is dependent upon the transportation of goods, personnel, and communications between geographically diverse locations. Commercial airlines, freight services, and internet-based auction sites all depend on access to efficient and effective transportation networks to maintain a competitive edge and continue profitable operations.

Profitability of civilian industry depends heavily on the agility and cost effectiveness of the transportation and supply chains used to transport raw materials for production and to ship finished goods to market. This need has led to the establishment and growth of Third Party Logistics providers (3PL). These 3PL organizations and DoD supply chain counterparts are under continuous pressure to provide increasingly efficient transportation options to their customers. This can be attributed to globalization, shrinking budgets, and expanding mission requirements. Meeting these demands requires distribution providers to integrate historically disparate transportation modes creating true multimodal options.

Military applications of transportation are equally dependent on the performance of the networks they utilize. The problem of transporting commodities, and people is so important to the DoD, that the DoD has attempted to centralize transportation control across the Army, Navy, and Air Force as early as 1979 with the formation of the Joint Deployment Agency [90]. In 1987 the United States Transportation Command (USTRANSCOM) was founded. Today the USTRANSCOM is responsible for all assets and commodities via the Surface Deployment and Distribution Command (SDDC), Military Sealift Command (MSC), and Air Mobility Command (AMC) until it reaches an Area of Responsibility (AOR) at which time responsibility for goods in transit is assumed by a Geographic Combatant Commander (GCC). In order to support these operations, USTRANSCOM must decide how to allocate the air, land (truck/rail), and sea assets as well as personnel resources in such a way as to ensure cargo is delivered on time to the locations it services [81]. The 2011 USTRANSCOM Strategic Plan discusses the need for JDDE synchronization across all modes of transportation utilized by the JDDE partners [107].

Although multimodal transportation networks are used every day in both military and civilian applications, the process of planning and coordinating transportation modes has received surprisingly little attention in the literature. Many recent publications in the field of Supply Chain Management and Transportation claim to address multimodal transportation planning but, in fact only address the coordination of at most two transportation modes [52]. Addressing multimodal transportation as an integrated network system is a relatively new concept. In fact the first instance of this systems approach found in the literature is from 1999 [88]. This research addresses capacity planning in multimodal freight networks.

1.2 Problem Description, Themes, & Gaps

Planning and coordinating multimodal transportation networks requires a method of estimating the ability of the network to meet demands. This estimate must explicitly consider many factors affecting the overall capacity that the network can support. Network capacity over a specified time period is sometimes referred to as throughput. This definition of throughput will be adopted in the remainder of this research. Freight network throughput is affected by many factors. Those considered in this research are: planning horizon length, fleet size for each transportation mode, payload (in tons) of each asset by type, asset utilization rate, block speed, and asset productivity.

Not only is the problem of estimating capacity in transportation networks difficult, but given the increased demands for network capacity due to increases in cargo volume demand, the problem has become a major concern of public policy [88].

Estimating the capacity of a multimodal network is plagued by the inadequacies in current capacity modeling methods [96]. Previous work in estimating the capacity of multimodal freight networks is inaccurate due to measuring individual system components or the capacity of partially aggregated components. These partial aggregations usually involve aggregation for estimating the capacity of single transportation modes or aggregation in estimating the capacity of single "corridors" through the multimodal system [49, 88, 105, 115]. The inadequacy stems from the inability of current models to capure interactions among transportation modes due to mid-path mode transfers and the interaction of other system components [96].

The most recent improvements in the area of multimodal freight transportation system capacity seek to provide a systems view of multimodal transportation networks but addresses the problem at a tactical level of detail and only provides a model formulation. No attempt is made to develop a solution approach for the bilevel programming problem resulting from the model formulation [96].

1.2.1 Themes. Three themes provide context for the existing literature and contributions: Integration of multiple fleets of transportation across multiple modes of transportation and mode/carrier selection. Modal integration addresses the problem of coordinating the resources of the different transportation modes in order to gain efficiencies through synergy. The literature review will demonstrate that modal integration is typically limited to two modes of transportation. Mode selection seeks the single "best" mode of transportation with respect to cost, security, delivery time, or some combination of these and other objectives.

Another emergent theme that is not addressed in the literature in the context of multimodal transportation is the identification of efficient or near-efficient solutions to capacity problems. An efficient or Pareto optimal solution is a multi-objective concept of optimality. An efficient solution is one which cannot be improved with respect to one objective without sacrificing quality in at least one other objective [40]. Efficiency will be discussed further in Chapter 5.

1.2.2 Gaps. In the field of production and inventory management, there is a keen sense of the fidelity required to execute planning during the long-term, mid-term, and short-term planning horizons. This awareness helps to control extraneous data collections which are often both costly and time-consuming. Data for high fidelity planning can be prohibitively difficult to obtain. Additionally, high fidelity planning (scheduling the production on a single machine over the course of an hour for instance), is not required in order to plan production operations over the course of the next year. The purpose of the long and mid-term planning is to estimate the feasibility of a desired level of production relative to the manufacturing system capacity. Long-term (aggregate) plans are refined via mid-term (rough-cut capacity) planning as the objective planning period draws closer. When the objective planning period is within a month, a master production schedule is created from the rough-cut estimates obtained through mid-term and long-term planning. In this way, the production capacity of the system is not overwhelmed in any period.

No current source in the literature provides a time scalable method of rough-cut multimodal freight transportation network capacity estimation. Scalability refers to the ability of the method to adapt capacity estimates to an arbitrarily defined planning horizon. Furthermore, the limited number of capacity estimation methods are based on tactical-level considerations and are not well suited to estimating overall system capacity without tactical-level data which may not be available [96]. This observation led to the identification of several gaps in the literature addressed by this research:

- (G1) No scalable multimodal freight transportation network capacity estimation method is found in the currently available literature.
- (G2) There is a lack of an operational-level multimodal freight transportation capacity estimation method requiring only operational-level data. Such a method would be an extension of "capacity planning based on overall factors" (CPOF) used in production and inventory management.
- (G3) Currently, there is no rough-cut capacity planning technique for multimodal transportation. Developing such a capability by extending existing concepts would represent a major contribution to the field of transportation research.
- (G4) Development of a new modeling approach that considers a multimodal transportation system holistically would represent a contribution to the field of transportation research.
- (G5) Development of a solution procedure to identify efficient solutions under various planning assumption scenarios would represent a contribution to the fields of operations research and transportation research.

1.3 Problem Statement

Identification of these gaps leads immediately to the following problem statement: "Given a fleet of vehicles from across multiple modes of transportation, demand quantities/locations, supply quantities/locations, and planning horizon length, determine the multimodal fleet mix and route providing the required capacity at minimum cost based only upon these operational-level data inputs"

The capacity of a network relates to its throughput over a specified period of time. Since network flow may be continuous, the amount of throughput must be defined for a specified time period in order to provide a basis for measurement and comparison. Capacity estimates in freight transportation networks are conveniently expressed in units of million ton miles per day (MTM/D). This standard measure of vehicle fleet performance is the standard for air mobility but may be calculated for any fleet of vehicles [91].

1.3.1 Research Questions. Effectively addressing this problem statement leads to several questions that must be answered. The specific questions of interest in this research are:

- (Q1.G1/G2/G3) Can concepts of Rough-Cut Capacity Planning be extended to estimate the capacity of a multimodal freight transportation network as a system given existing routes?
- (Q2.G2/G3) Can modeling and solution techniques be developed to allocate the available capacity of a multimodal freight transportation network in a way that minimizes costs and meets demand timelines?
- (Q3.G2/G3) What multimodal routes should be established in a network, in order to improve cost efficiency and mission effectiveness?
- (Q4.G5) How do changes in numbers of assets, available routes, planning horizon length, and input costs (fuel for instance) affect mode choices and overall system costs?
- (Q5.G4) What areas of the multimodal transportation system would benefit most from resource investments?
- (Q6.G1/G2) What is the appropriate fleet mix required to fulfill demand in a set planning horizon at a minimum cost to the overall system?

- (Q7.G5) How *robust* is the multimodal transportation system to losses in established routes or assets? That is, what is the nature and length of system capability degradation due to these losses?

1.4 Research Objectives

The hypothesis of this research is that multimodal rough-cut system capacity planning can be executed using a network-based approach and capacity planning techniques normally applied to production and inventory management. Extending these methods requires the development of new theory. The following research objectives provide original theoretical contributions and move the state-of-the-art forward with respect to closing the gaps identified above.

Addressing these questions requires the development of a new methodology and new modeling and analysis approaches to augment what is currently available in the literature. These questions lead to an explicit statement of the research objectives.

1.4.1 Research Objectives & Contributions.

- (O1.Q1/Q2/Q3) Develop a methodology for operational-level multimodal freight transportation capacity estimation that is scalable to an arbitrary planing horizon and an arbitrary multimodal fleet size. The modeling method should capture interaction between the various modes of transportation.
- (O2.Q4/Q5/Q6) Develop a method for determining the least cost fleet mix and routing structure for a specified range of assets and a specified range on planning horizon.
- (O3.Q7) Develop a method for generating alternative near efficient solutions and solutions with passive resilience to aid decision makers in analyzing tradeoffs in time, cost, and fleet assets.

1.5 Dissertation Overview

Chapter 1 has motivated this dissertation effort by briefly highlighting the nature of the problem, discussing the relevance of the research within the context of both military and commercial applications, and has presented the intended contributions.

Chapter 2 will establish the uniqueness of the research in this dissertation by providing a thorough treatment of the relevant literature. This chapter will compare and contrast this dissertation work with past work. This chapter shows that the intended contributions are indeed original and provide a solid scope and context for the Methodology.

In chapter 3 we show how Multimodal Rough-Cut Capacity Planning is modeled using the Resource Constrained Shortest Path Problem. We demonstrate how this approach supports either mode selection or integration depending upon transportation mode costs, fleet size, and planning horizon.

In chapter 4 we develop an Ant Colony System metaheuristic to quickly solve large instances of the RCSP. In chapter 5 we develop a heuristic based on the ACS of chapter 4. This new metaheuristic generates near-efficient solutions and helps identify passive resilient solutions to the multimodal RCCP problem to aid decision makers in analyzing tradeoffs between time, cost and asset utilization.

Chapter 6 summarizes the contributions, results, and conclusions of the dissertation and presents final analyses. Recommendations for future work complete this chapter.

II. Literature Review

2.1 Literature Review Strategy

This chapter provides a broad overview of the relevant literature. Chapters 3, 4, and 5 contain context specific literature review and thus the purpose of this chapter is to demonstrate the uniqueness of the dissertation research described in Chapter 1.

2.2 Overview of the Literature

Figures 2.1 and 2.2 provide a quick reference summary of the important sources reviewed in this dissertation. Some sources, such as textbooks, were omitted since their purpose is to support the analysis rather than contribute original work to any field of research.

	Primary Motivation									Methodology Areas													
	Vehicle Routing/Scheduling	Traveling Salesman Problem	Transportation Mode Selection	Transportation Mode Integration	Transportation Fleet Capacity Plannir	Vehicle Loading/Pallet Packing	Transportation Network Design	Heirarchical Production Planning	Network Reslilence	Linear Programming	Binary/Integer Programming	Bi-Level Programming	Multi-Objective Optimization	Scheduling	Tabu Search Metaheuristics	Ant Colony Metaheuristics	Heuristics/Approximation Other	Solving the RCPSP	Solving the RCSP problem	Goal Programmin	Networkk Contraction/Aggregation	Deterministic Algorithms	Solving k-shortest paths problem
Dissertation Research			•	•	•		•		•		•		•			•			•				
Amiouny, Bartholdi, and Vande Vate						•											•						
Andersen, Crainic, and Christiansen							•				•			•									
Asakura					•					•													
Avella, Boccia, and Sforza (1)					•						•								•		•		
Avella, Boccia, and Sforza (2)																	•		•				
Ballew						•					•						•						
Balseiro, Loiseau, and Ramonet	•															•	•						
Bard and Moore										•		•										•	
Barnhart and Ratliff				•																		•	
Barton										•						•							
Barton and Hearn							•			•	•						•				•		
Beasley and Christofides											•								•				
Bitran, Haas, and Hax								•		•				•									
Bodin and Golden	•																						
Boland, Dethridge, and Dumitrescu																			•			•	
Bonabeau, Dorigo, and Theraulaz																•	•						
Calhoun	•									•	•			•						•			
Coello, Coello, Pulido and Lechuga													•				•						
Derringer, and Suich													•										
Dongarra										•	•												
Dorigo; Dorigo & Gambardella; Dorigo & Stutzle; Dorigo, Birattari, & Stutzle; Dorigo, Maniezzo, & Colorni		•														•							
Dumitrescu and Boland										•							•		•			•	
Eppstein																							•
Elimam and Kohler											•							•					
Feillet, Dejax, Gendreau, and Gueguen	•																		•			•	
Gambardella and Dorigo		•														•							
Gan, Guo, Chang, and Yi		•														•							
Garcia																	•		•			•	

Figure 2.1: Overview of Reviewed Literature Sources

	Primary Motivation									Methodology Areas													
	Vehicle Routing/Scheduling	Traveling Salesman Problem	Transportation Mode Selection	Transportation Mode Integration	Transportation Fleet Capacity Planning	Vehicle Loading/Pallet Packing	Transportation Network Design	Heirarchical Production Planning	Network Resilience	Linear Programming	Binary/Integer Programming	Bi-Level Programming	Multi-Objective Optimization	Scheduling	Tabu Search Metaheuristics	Ant Colony Metaheuristics	Heuristics/Approximation Other	Solving the RCPSP	Solving the RCSP problem	Goal Programmin	Networkk Contraction/Aggregation	Deterministic Algorithms	Solving k-shortest paths problem
Dissertation Research			•	•	•		•		•		•		•			•			•				
Ge, Zhang, and Lam									•			•					•						
Handler and Zang											•								•			•	
Hassin																	•		•				
Hoffman and Pavley																							•
Hu, Liu, and Liu.	•															•			•				
Irnich and Desaulniers	•										•						•		•			•	•
Irnich and Villeneuve	•												•						•				
Ji, Hu, Zhang, and Liu							•									•							
Kaluzny and Shaw						•				•	•											•	
Katoh, Ibaraki and Mine.																						•	•
Kjerrstrom					•						•								•				
Lorenz and Raz																	•		•				
Macharis and Bontekoning	•		•	•	•	•	•																
Martins						•					•						•					•	
Mehlhorn and Ziegelmann										•									•			•	
Meixell and Norbis			•																				
Merkle, Middendorf, and Schmeck																•		•					
Moore and Bard										•		•										•	
Morlok and Chang.					•				•			•											
Morlok and Riddle					•					•											•	•	
Mullen, Monekosso, Barman, and Remagnino																•							
Nasiri																	•	•					
Park and Regan			•									•									•	•	
Ribeiro and Minoux																	•		•				
Sun, Turnquist, and Nozick					•				•		•												
Unnikrishnan and Waller					•				•	•													
Van Hove	•										•			•			•	•				•	
White	•															•							
Yagmahan and Yenisey													•	•		•							
Yang, Bell and Meng					•				•			•					•						
Zhang, Li, and Tam																	•	•					
Zhu and Wilhelm											•								•			•	
Zitzler and Thiele													•				•						
Zografos and Regan	•		•	•	•			•															

Figure 2.2: Overview of Reviewed Literature Sources

Literature	Mode Selection and Integration Capabilities	Multi-Objective Analysis of Time/-Cost/Asset Tradeoff	Generate alternative near-efficient and passive resilient solutions
Math Programming	x	x	x
Multi-Objective		x	
Scheduling			
Heuristics		x	

Table 2.1: Features of this dissertation versus literature's research

Capacity planning in transportation is usually limited to one or two modes of transportation. Sources considering more than two modes of transportation are considering the problem of mode selection [83, 96]. True multimodal integration is a topic which is considered an area of opportunity in recent surveys of the transportation field [79, 119]. The focus of this dissertation is on the development of methods for capacity analysis which are scalable to an arbitrary number of transportation modes. Table 2.1 provides a concise comparison of this dissertation with other areas of reviewed literature.

Sources in the literature focus on either mode selection or mode integration. This dissertation is unique in several ways including the capability of analyzing a multimodal network considering both mode selection and mode integration. A single mode solution is returned if it is the most cost effective or is "best" with respect to the multiple defined objectives. A multimodal solution is returned if it is the most cost effective, is the best with respect to the multiple objectives, or if multiple modes of transportation are needed in order to make the problem feasible by providing the required capacity to meet delivery demands in the specified planning horizon. Table

Chapter	Unique Transportation Planning Capability	Ops Research Methodology
3	Multimodal RCCP	RCSP Problem
4	Accessibility of Large Multimodal RCCP Analyses	Ant Colony System (ACS)
5	Generate Alternative Near-Efficient and Passive Resilient Solutions	New Heuristic based on ACS

Table 2.2: Overview of this dissertation

2.2 presents a summary of the content of the dissertation development in chapters 3, 4, and 5.

Capacity planning can be viewed from either the aggregate or the rough-cut perspective. Aggregate planning focuses on the longer term strategy of production and attempts to ensure supply matches demand by prescribing monthly or weekly production output levels. By contrast, rough-cut planning seeks to ensure that the production level prescribed in the aggregate planning process is feasible by estimating the potential throughput or capacity of the production system.

2.3 Multimodal Freight Transportation

Literature sources reviewed in this research use the terms *intermodal* and *multimodal* interchangeably. In order to standardize their use in this dissertation the terms shall be defined as follows:

- *Multimodal* transportation refers to the delivery of freight from origin to destination involving at least two different modes of transportation.

- *Intermodal* transportation is refers to the transfer operations to move freight between two modes of transportation at container terminals or other types of intermodal ports [23].

2.3.1 Intermodal Transportation. *Intermodal transportation* is defined as: "the transportation of a person or load from its origin to its destination by a sequence of at least two transportation modes, the transfer from one mode to the next being performed at an intermodal terminal" [29]. The process of moving the goods between modes is often referred to as *transshipment*. It is not a goal of the research which follows to treat transshipment with any level of detail. Rather, this section is designed to acknowledge that this topic is indeed an area of research in itself. For further review of the topic of transshipment the reader is referred to [9, 76, 85, 108, 112]. These efforts are targeted at the tactical level and focus on port operations including how to utilize different technologies to efficiently accomplish port tasks. For the purposes of this research, tactical efficiencies matter only as they relate to throughput capacities of the various nodes (bases) in the network.

A 1996 article by Haghani and Sei-Chang models logistics for disaster relief as a multi-commodity, multi-modal network flow problem [53]. The modeling approach utilizes the concept of a time-space network which has also been used by Nielsen et al. and in many other applications by Barnhart and Armacost in a similar manner [6, 7, 95]. The authors formulate the network as a MILP and then propose two different heuristics for obtaining solutions to the formulation. Route length restrictions are not considered in this formulation. By using the time-space network formulation, the model is capable of representing a change in transportation mode by an arc traversal. Similarly, advances in time are represented as arc traversals as well. The distinction between changes in mode, changes in time, and changes in location are managed by indexing the sets of time-windows, locations, and modes.

2.4 Multi-Layered Networks

Kennedy introduced the concept of multi-layered network synthesis considering both cost and robustness as competing objectives. He applied a multicriteria decision analysis technique using epsilon constraints to scalarize the objective function. He then extended this new formulation to include flow and connectivity constraints [73]. Multimodal freight transportation networks can be modeled using layered networks and interactions (i.e. transfers between modes) can be modeled using inter-layer arcs [73, 74].

2.5 Hierarchical Production Planning

Hierarchical Production Planning techniques have been developed for use in planning production operations at varying levels of fidelity. Long-range planning is generally executed using lower levels of fidelity since there may be changes in important factors affecting production between the planning and execution phases. Mid-range planning is somewhat higher fidelity but still does not produce detailed tactical level schedules due to potential changes in input factors. Master Production Scheduling produces detailed schedules of how the raw inputs will be converted into finished products using the manufacturing resources of the factory, like machines and manpower [19, 30, 58, 60, 61, 69, 93, 98].

2.5.1 Aggregate Production Planning. Aggregate planning encompasses both aggregate output planning and aggregate capacity planning. Aggregate output planning attempts to match production and demand. Aggregate capacity planning assesses the feasibility of the aggregate output plan by estimating how much production capacity is required to meet the production levels prescribed by the aggregate output plan [69]. In aggregate planning the individual products are normally aggregated into product families based on some measure of product similarity. By reducing the number of products the aggregate planning process can be completed in a timely manner [28]. The estimates obtained using aggregate planning are rough, but are usually sufficient to estimate the long-term requirements in the six to eighteen month rolling planning horizon.

2.5.2 Rough-Cut Capacity Planning. The focus of this dissertation is rough-cut capacity planning in multimodal transportation. Rough-cut capacity planning is similar to aggregate capacity planning but differs in fidelity and function. In rough-cut planning the goal is to estimate the ability of the system to meet the required production demand over a shorter time period and is used to assess the feasibility of a master production schedule. RCCP addresses the question: "Is there sufficient capacity in the production capabilities to support the desired level of production?" There are three prominent methods of performing RCCP: Capacity Planning using Overall Factors (CPOF), Capacity Bill Procedures, Resource Profiles [66, 109].

In the context of this dissertation, rough-cut capacity planning will be defined as planning to establish the feasibility of a desired level of multimodal system freight throughput in millions of tons. The scheduling of the multimodal fleet is not addressed but many heuristics for scheduling fleets exist and the reader is encouraged to reference Pinedo for a thorough treatment of scheduling algorithms and heuristics [97].

The function of the capacity planning and the data requirements are the primary reasons for classifying this dissertation as rough-cut capacity planning rather than aggregate capacity planning. As Chapter 3 will show, only operational level data such as aircraft capacity, and speed are needed in order to complete the capacity planning approach developed in this research. Also, the heuristic developed is meant to provide a user with the ability to quickly produce solutions under different scenarios so the function is more in line with near-term rather than long-term planning. However, the math programming model provided in chapter 3 may also be solved by deterministic methods if time allows so its scope is not restricted only to near term estimates as optimal solutions may be determined as well.

2.5.3 Master Production Scheduling. The purpose of master production scheduling is to plan the short-term (i.e., weekly or monthly) production activities of the factory in order to ensure demand is met. In MPS the product families are disaggregated into individual products and are treated individually. A key requirement for the successful linking of aggregate planning, rough-cut planning, and master production scheduling is in aggregating and disaggregating products and product families in an appropriate way. For more detail on the issues associated with aggregation and disaggregation the reader is encouraged to reference Hax and Candea [58].

2.6 Capacity Planning in Multimodal Freight Transportation

Park developed a first-of-its-kind model to address multimodal freight transportation from a comprehensive standpoint in order to leverage the competitive advantages of each mode [96]. The model Park develops is an integrated multimodal freight model but the model uses an "average load factor" for each mode and therefore is not additive with respect to capacity over the entire fleet of vehicles in each mode. This is in contrast to the MTM/D approach proposed in this research [96].

Sun et.al. use an existing LP model called MAXCAP which seeks to maximize the system capacity for multiple origin destination pairs subject to resource and capacity constraints. Additionally, they analyze the "flexibility" and performance of the system under degradation. However, they consider only a single mode of transportation, movement by rail [104].

Andersen et. al. address the design and optimization of a multimodal scheduled service network. The model provides "tactical" level planning. The main goal of the model developed is to determine the appropriate departure times to maximize throughput. The primary issue in this report is the coordination and synchronization of schedules among multiple fleets of vehicles [4].

Kasikitwiwat and Chen discuss different definitions for network capacity. The authors categorize capacity models into two main types: economic and physical network capacity [71]. Three methods of examining physical network capacity (reserve, ultimate, and practical capacities) are presented and discussed.

Russ et.al. formulate a network design problem for multimodal transportation infrastructure investment planning in Indonesia, as a bi-level programming problem. The lower-level program models traffic flow equilibrium and is solved using a iterative descent direction algorithm. The upper level program manages the capacity allocation on the arcs and is solved using a genetic algorithm [102].

2.7 Shortest Path and Resource Constrained Shortest Path Problems

The RCSP is a generalization of the well known shortest path problem in which the elements of the shortest path are also subject to knapsack constraints based upon resource limitations. The RCSP was first proven to be NP-Complete in 1980 by Handler and Zang and many extensions and applications have since been developed [38, 48, 54]. Even the least general version of the RCSP, the Elementary RCSP (ERCSP) in which no arc is repeated has been shown to be strongly NP-Hard by Beasley and Christofides in 1989 [18]. This ERCSP finds applications in vehicle routing where it is often the route pricing subproblem [26, 51, 101].

III. Rough-Cut Capacity Planning using Resource-Constrained Shortest Paths

3.1 Introduction

Consider the problem of planning for capacity utilization in multimodal freight distribution networks. This complex problem has received relatively little attention in the literature from a truly integrated multimodal perspective [79]. Capacity allocation in multimodal freight transportation must address many factors in determining to which routes capacity is to be allocated.

The literature sources reviewed demonstrate a gap in the area of transportation mode integration. Much of the previous research addressing multiple transportation modes focuses on mode selection rather than integration. Mode integration is defined here as the seamless use of two or more modes of transportation to provide required system capacity. Mode integration differs from mode selection in that mode selection is concerned with what percentage of demand is to be allocated to each mode of transportation, usually based upon cost. These decision problems are typical in third party logistics (3PL) applications. A weakness of this approach is that mode selection views transportation modes in a strictly parallel configuration. This view of the transportation modes as separate systems ignores the possibility of cooperation among multiple modes to save time and money.

A recent survey of the transportation mode choice and carrier selection literature reviewed over ninety separate references spanning a period of over forty years [83]. These efforts seek the mode that provides the least expensive shipping option to support delivery of the required freight capacity within the specified time period. Research efforts following this methodology ignore the potential for even greater cost and time savings that might be gained through truly integrated multimodal shipping options.

Many articles have been devoted to estimating the capacity of single mode transportation networks and, in particular, rail network capacity has been extensively studied. Many of these efforts use or extend the MAXCAP model which uses "traffic lanes" that are aggregated network nodes geographically close with an established recurring commodity flow [87,88,104]. Some "multimodal" capacity planning sources surveyed were, in fact, only bi-modal [16]. This is an obvious inadequacy in the literature since many logistics providers use a combination of air, truck, rail, and sea assets. Methods of simultaneous mode analysis are badly needed.

Previous network capacity modeling efforts are inadequate to capture the true transportation network capacity since they typically utilize network aggregation techniques which reduce the fidelity and accuracy of the resulting capacity estimate. *Corridors* and *modal networks* are two popular methods of aggregation used in estimating single and multimodal network capacities [49,88,105]. A recent modeling effort focusing on multimodal transportation capacity estimation uses a bi-level programming formulation to allocate capacity. The formulation is so complex no attempt is made to provide a solution method [96]. Aggregation is done *a priori* resulting in a network abstraction which has no ability to yield insights at higher fidelity. The final result is an inadequate analysis resulting from a loss of valuable system information due to aggregation.

Macharis and Bontekoning have established opportunities in operations research for contributions to multimodal transportation [79]. The gaps in the literature are categorized into drayage, terminal, network, and intermodal categories based upon the level of problem fidelity. This work considers the "network" and "intermodal" areas. *Drayage* is mainly used in reference to the transport of containerized cargo over short distances. *Terminal* operations are the port operations concerned with transloading cargo between modes or vessels of conveyance *Network* issues concern infrastructure and route planning and pricing. *Intermodal* or *multimodal* issues surround the route and service choices for existing networks. In particular there are opportunities in methods that aid intermodal operators in determining which group of routes and services to purchase from terminal and drayage operators on behalf of a shipper to minimize cost and delivery time frame.

Zografos and Regan discuss challenges in intermodal freight transportation in Europe and the United States and the need to develop information and communication technologies to aid decision making in a deregulated, free enterprise transportation market [119]. They reiterate the need for planning methods to aid 3PL firms in purchasing individual services from separate providers that can act as integrated multimodal transportation options.

The "Multimodal Corridor and Capacity Analysis Manual" suggests the use of multimodal corridors for which capacities are calculated in aggregate [105]. One inadequacy of this approach is that any change or disruption in a link within the aggregated corridor capacity estimate would require a separate capacity analysis.

Other network capacity estimation research efforts lack the ability to model multimodal freight transport capacity and allocation. Asakura estimates the capacity of urban transportation networks given user route preferences using a bi-level programming formulation [8]. This bi-level programming approach is also used by Park and Regan to estimate the capacity of multimodal transportation systems although the model is not solved due to the complexity of the formulation [96]. Yang et al. formulate a similar network capacity problem which further accounts for the potential for land development in determining route selection. Once again this approach relies on a bi-level programming formulation for which a heuristic is proposed [115]. Morlok and Riddle rely upon the MAXCAP problem formulation to estimate transportation system capacity. This model uses aggregated routes based on *a priori* information of commodity flow between origin-destination pairs [88]. Ge et al. use a bi-level programming formulation to determine the effect of reserve capacity on a road network while accounting for traveler information [49]. Methods not relying upon bi-level programming to plan system capacity are preferred since bi-level programming problems require development of specialized algorithms. In the case of mixed-integer bi-level programming problems solving the relaxation does not provide a valid bound on the optimal solution to the original problem causing many difficulties in even obtaining feasible solutions to capacity problems formulated as bi-level programming problems [15, 86].

The preceding paragraphs provide strong evidence that simply extending current modeling techniques is not sufficient to overcome the inherent inadequacies in current capacity modeling approaches. Therefore, a new capacity modeling methodology is required in order to improve the fidelity and accuracy of network capacity estimates. In the context of freight transportation, such an approach must capture important system parameters like single trip vehicle capacity, length of planning horizon and the size of the vehicle fleet. These parameters, and others, drive network capacity. Allocation of this capacity should provide network paths efficiently using fleet capacity.

Minimizing the data collection required is highly desirable since data collection is often difficult, time-consuming, and expensive. Reusing data that are already available eliminates the need to collect and analyze additional data sources which may support no other purpose. This article explicitly addresses all of these factors and develops a new capacity modeling approach for capacity planning. Our novel approach is then extended in order to model capacity given an arbitrary number of distinct transportation modes.

Given a fleet of vehicles from across multiple distinct modes of transportation, single source and sink, and planning horizon length, consider the problem of determining the multimodal fleet mix and path which will provide the required capacity at minimum cost based only upon operational-level data inputs.

We define operational data as those data that are available without requiring additional data collection. Examples would be engineering and operating specifications of the vehicle fleets, estimates (in tons) of the freight to be transported, and the length of the time window for transportation planning.

We demonstrate how rough-cut capacity planning (RCCP) in freight networks is modeled using a modification of the Resource Constrained Shortest Path Problem (RCSP). The basic Shortest Path Problem (SPP) is modified by adding a knapsack constraint on resource consumption. The RCSP was originally proposed by Handler and Zang in 1980 [54]. Define the following sets and variables:

$$
\begin{array}{ll}
N = \{1, 2, ..., n\} & \text{a set of uniquely labeled nodes} \\
A = \{(i, j) : i, j \in N\} & \text{a set of directed arcs that define} \\
 & \text{adjacencies for the nodes in } N \\
G = (N, A) & \text{is a graph defined by the nodes in} \\
 & N \text{ and the arcs in } A \\
R & \text{is the quantity of resource available} \\
s & \text{is the label assigned to the source} \\
 & \text{node where supply is located} \\
t & \text{is the label assigned to the destination} \\
 & \text{node where demand is located} \\
c_{ij} & \text{is the cost of traversing arc } (i, j) \\
x_{ij} = & 1 \text{ if arc } (i, j) \text{ is included in the path} \\
 & \text{and 0 otherwise}
\end{array}
\tag{3.1}
$$

The RCSP is formulated as:

$$
\begin{align}
\text{Minimize} \sum_{(i,j)\in A} c_{ij}x_{ij} \tag{3.2a}\\
\text{s.t.} \sum_{j:(i,j)\in A} x_{ij} - \sum_{j:(j,i)\in A} x_{ji} = b(i) \qquad \forall i \in N \tag{3.2b}\\
\sum_{(i,j)\in A} c_{ij}x_{ij} \leq R \tag{3.2c}\\
b(s) = 1, b(t) = -1, b(i) = 0 \ \forall\ i \neq s, t \tag{3.2d}
\end{align}
$$

In this formulation, 4.2a is the objective function that minimizes cost, 4.2b are the flow balance constraints, 4.2c are the knapsack constraints on resources, and 4.2d are the flow forcing constraints.

Applications of the RCSP to transportation include solving vehicle routing problems [43], routing a homogeneous vehicle fleet [11], in such diverse contexts as wastewater treatment in Kuwait and in ensuring compliance with government standards for thermal resistance in building design [41].

The RCSP is at least as hard as NP-complete problems as first shown by Handler and Zang in 1980 [5, 48, 54, 75]. RCSP's can be classified according to four basic criteria [64]: resource accumulation, path structural constraints, objective function, and the nature of the underlying network.

Resources may be accumulated either on the arcs or at the nodes. Accumulation of resources on the arcs uses *resource vectors* while accumulation at the nodes uses *resource intervals*. These two methods for resource accumulation are equivalent if there is no lower bound on resource consumption. Path structural constraints capture node-arc adjacencies, branch-and-price rules, and admittance of any non-simple paths to the solution. Objective functions are classified as single-objective or multi-objective depending upon the number of objective components being considered. Finally, the underlying network is classified as cyclic or acyclic. Networks containing cycles have an infinite number of paths and the solutions may be unbounded. However, cyclic networks can be transformed into acyclic networks via discretization to prevent unbounded solutions [64].

We now turn our attention from classifying RCSP problems to classifying RCSP solution methods. Garcia, and Mehlhorn and Ziegelmann have provided classification schemes for RCSP's [47,82] which are broadly categorized into path ranking, labeling, and Lagrangian Relaxation methods.

Path ranking methods are based upon the work of Eppstein, and of Hoffman and Paveley who first studied the k-shortest path problem [42,59,72]. Such methods generally operate by building paths and then ordering them according to some criteria such as cost or length. In RCSP's the shortest path often breaks one or more of the knapsack constraints. Selecting an alternative path (from the rank ordered k-shortest generated) which is also feasible in the knapsack constraints is one approach using k-shortest paths.

Labeling methods operate by searching for non-dominated paths. Each path is labeled with both its cost and resource consumption. When an alternate path is found that is superior to the incumbent with respect to either cost or resource consumption and at least as good in the remaining criterion, then the incumbent is replaced by the alternate path. Irnich and Villeneuve provide a labeling method for k-cycle elimination for $k \geq 3$ [65].

Lagrangian Relaxation-based solution methods solve a relaxation of the original problem by relaxing the knapsack constraint associated with the resource consumption, and then working to close the duality gap between the relaxed linear programming primal and dual solutions. Mehlhorn and Ziegelmann have created a RCSP software package called "Constrained Network Optimization Package" which operates based on Lagrangian Relaxation and closing the primal-dual gap [11,18,21,100,113].

Heuristics other than Lagrangian Relaxation have also been successfully applied to the RCSP and similar problems. Particle Swarm Optimization was applied to the Resource Constrained Project Scheduling Problem by Zhang et. al. [116]. Hassin developed a fully polynomial approximation bounding scheme for the RCSP that was later generalized by Lorenz and Raz to accommodate both positive and negative arc weights [57, 78]. Hu et al. develop an improved version of the Ant System (AS) originally proposed by Dorigo for solving the RCSP [63]. The proposed AS relies upon modified heuristic information and pheromone updates to drive convergence to paths meeting the resource constraints [33].

Zhu and Wilhelm developed a three step approximation algorithm that is applied iteratively to obtain successively tighter bounds on the optimal solution [117]. Step one preprocesses the graph to remove nodes and arcs which cannot be contained on any feasible path. Step two expands the reduced graph based upon resource availability and step three solves the unconstrained shortest path problem. Arc weights are adjusted and the algorithm is repeated until sufficiently tight bounds are determined. There are many other research efforts that have been devoted to developing fast running heuristics for the RCSP [10, 39, 57, 75].

Jepsen et al. provided the first solution method for the ERSCP based on branch-and-cut. In order to strengthen the formulation a new class of valid inequality called *generalized capacity inequalities*, is developed. This new class of inequality is combined with knapsack constraints and subtour elimination constraints which are all introduced as cutting planes in the branch-and-cut algorithm [67].

No source reviewed in the literature provides a time scalable method of multimodal freight transportation network capacity estimation. Scalability refers to the ability of the method to adapt capacity estimates to an arbitrarily defined planning horizon. Furthermore, the limited number of capacity estimation methods are based on tactical level considerations and are not well suited to estimating overall system capacity without tactical level data which may not be available [96].

3.1.1 Assumptions. For the purposes of this research we assume that capacity is expressed in terms of tons. This assumption allows us to use million-ton miles (MTM) as the resource for the knapsack constraints as we demonstrate in the next section.

We do not consider weight and balance, dimensional packing, or individual vehicle routing as these problems are well represented in the literature [3,13,20,70,80].

Finally, we assume that scheduling is completed in near-term planning. RCCP is targeted at mid-range planning and answering the question "Does the system have sufficient capacity to meet the demand." Once demand and capacity are equilibrated through RCCP, scheduling algorithms and heuristics are used to build final delivery schedules [58, 97].

3.2 Modeling Approach

The modeling approach established in this section treats each transportation mode as a separate layer of a network. Each network layer shares a common node set with all other network layers. The layers differ by the arc sets which define node adjacencies within each layer. In order to ensure that the available capacity within each mode of transportation is respected by the modeling approach, the RCSP is used to capture the capacity constraints which are defined for each layer. These capacity constraints will be defined in the following sections.

The RCSP is preferred to the Multicommodity Network Flow Problem (MCNF) presented by Ahuja et al. since the MCNF uses bundling constraints on total arc flow shared by all commodities and cannot capture individual capacity constraints for the various modes of transportation [2]. By contrast, the knapsack constraints on resources can be used to constrain resource usage through summing arc weights along a network path.

Modeling the rough-cut capacity problem for multimodal transportation requires the development of some model parameters. First, we formulate the single mode case as a RCSP problem and then generalize the result to an arbitrary number of distinct transportation modes. In a physical network, each pairwise adjacent set of nodes has associated with it, a distance label. The graph is not required to be complete. That is, not all nodes in the network are pairwise adjacent. For example, consider a network of four shipping and receiving locations. Let node 1 be a source node, node 4 be a sink node, and nodes 2 and 3 be transshipment nodes. The arc labels represent the Euclidean distance between nodes. Such a network is shown in Figure 3.1:

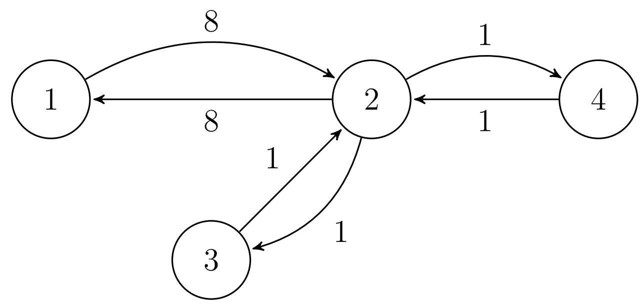


Figure 3.1: An Example Network: G^1

Network capacity can be viewed as the ability of a network to provide a required level of throughput within a given time window. Since network flow may be continuous, the amount of throughput must be defined for a specified time period in order to provide a basis for measurement and comparison. Capacity estimates in freight transportation networks are conveniently expressed in units of million ton miles per day (MTM/D). This standard measure of vehicle fleet performance is the standard for air mobility but may be calculated for any fleet of vehicles [91]. Implicit in this definition of capacity is that the "cost" associated with adding capacity to a route, concerns distance, time, and required capacity (expressed in millions of tons).

In figure 3.1, for example, the distance between node 1 and node 2, $d_{12} = d_{21} = 8$. If five million tons of freight must be transported between nodes 1 and 2 at a cost of \$2 per MTM then we can re-weight the arcs in the network, expressing weights as the cost for the required million-ton-miles (MTM) allocated to an arc by multiplying each d_{ij} by the number of tons of freight (in millions) required to be moved, m, and the cost per MTM h. This label must then be weighted by the estimated cost per MTM for the mode of transportation. The final arc weight is then given by: $cs_{ij} = d_{ij} * m * h$. The arc weights now express a composite quantity which addresses cost, distance, and required capacity. In order to assign capacity to arc (i, j), c_{ij}/h MTM are required. The reweighted network is labeled as shown in Figure 3.2:

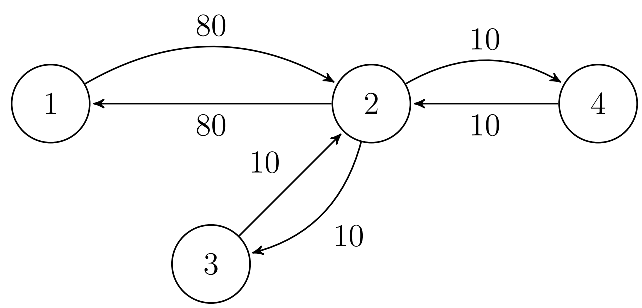


Figure 3.2: Reweighted Network

The number of MTM available, r, is a function of the transportation fleet characteristics and the planning horizon length (T). This can be expressed as: $r = V * MT * T$ where MT is the million-ton-mile per day capability of a single vehicle and V is the number of vehicles in the fleet. By inspection the shortest path through the example network is $1 \to 2 \to 4$. This path requires 45 MTM at a cost of \$90. If our fleet cannot provide the required MTM, then the problem is infeasible unless we add assets to the fleet (increase V) or add days to the extend the planning horizon length (T).

The parameter MT is calculated based on the following operational level data: "UTE" rate, payload, block speed, and productivity. UTE rate is the average historical "surge" capability of a single vehicle in hours per day. This represents the maximum availability of each vehicle per day. Down time due to scheduled maintenance and reliability issues are assumed to be captured in this quantity. Vehicle payload is the single trip capacity of a vehicle in tons. Vehicle block speed is the average door-to-door vehicle speed. Productivity, refers to the average historical percentage of non-empty trips over all vehicles. This parameter accounts for repositioning of vehicles within the fleet when they are moved while empty in order to pick up a load of freight. Now MT is calculated as $MT = MTM/D$ per asset $=$ ($UTE*$Payload$*$Block Speed$*$Productivity)$/1,000,000$. The single mode formulation can now be presented building upon the development of these parameters.

3.2.1 Single Mode Formulation. To simplify the presentation of the single mode resource constrained shortest path formulation the following parameters and vectors are defined: $N = \{1, 2, ..., n\}$ is the set of nodes representing shipping locations, $C = \{(l, v) : (l, v) \text{ is an arc in the graph}\}$ is the set of feasible directed arcs defined over the elements in N, $G = (N, C)$ is the graph consisting of nodes in N and arcs in C, and E is the $n \times |C|$ node-arc adjacency matrix where each column of E has a 1 in the l^{th} row and a -1 in the v^{th} row corresponding to adjacency matrix element E_{lv}. Now define E_{lv} along with the following sets, matrices, and scalars as:

$$E_{lj} = \begin{cases} 1 & \text{if } l \in N \text{ is the tail of arc } j \in C \\ -1 & \text{if } l \in N \text{ is the head of arc } j \in C \\ 0 & \text{otherwise} \end{cases} \quad \forall \quad l \in N,\, j \in C$$

$$q\_s(j) = \begin{cases} 1 & \text{if } E_{sj} = 1 \\ 0 & \text{otherwise} \end{cases} \quad \forall \quad j \in C$$

$$q\_t(j) = \begin{cases} 1 & \text{if } E_{tj} = -1 \\ 0 & \text{otherwise} \end{cases} \quad \forall \quad j \in C$$

$$Q = \begin{bmatrix} q\_s \\ \ldots \\ q\_t \end{bmatrix} \tag{3.3}$$

$h :$ the cost (in dollars) per MTM

$r :$ the number of MTM available

$B\_s :$ the $1 \times |C|$ vector of arc weights

$x :$ the $1 \times |C|$ vector of decision variables

$$J = \frac{B\_s}{h}$$

The single mode formulation can now be expressed as:

$$\text{Minimize} B\_s\ x \tag{3.4a}$$

$$\text{subject to:} Ex = 0 \tag{3.4b}$$

$$Qx = 1 \tag{3.4c}$$

$$Jx \leq r \tag{3.4d}$$

$$x \in \mathbf{B}^{|C|} \tag{3.4e}$$

By comparison with the generic RCSP formulation presented in section 3.1, this formulation is also a RCSP problem where 5.2a is the objective function, 5.2b are the transshipment portion of the flow balance constraints, 5.2c are the source and sink portion of the flow balance constraints, 5.2d are the knapsack constraints, and 5.2e are the flow forcing constraints.

3.2.2 Multimodal Formulation. Generalizing the single mode model to incorporate multiple transportation modes requires introduction of several other sets and indices to aid in describing not only the intra-mode arcs for each transportation mode but also the inter-mode arcs. To develop the generalization, an identical "copy" of the physical node set is created for each mode of transportation. Although the nodes for each mode of transportation are identical, the edge set for each mode of transportation may be distinct in order to represent different node adjacencies that might exist depending upon the mode of transportation. Furthermore, some nodes may act as intermodal transshipment points at which freight may be transfered between modes of transportation. A port which has both a rail head and a sea port may be one such combination. These intermodal transfers require definition of new sets of intermodal arcs. An example is provided following the development of the multimodal RCSP formulation.

We first define the following sets and parameters to generalize the single mode for addressing multimodal freight transportation. Let $P = \{1, 2, \ldots, p\}$ be the set of available transportation modes and $k \in P$ be be the index for these modes. Also $N = \{1, 2, \ldots, n\}$ be the set of nodes, as in the single-mode formulation with $i \in N$ being the index for the nodes. Expanding the original graph as described in the previous paragraph results in a node set with np nodes. The indices for the node set of the expanded graph are $l, v \in \{1, \ldots, np\}$. Let $N^k = \{1+n(k-1), \ldots, n+n(k-1)\}$ be the node set for mode k in the expanded graph. Finally, let $M^i = \{i, \ldots, i+n(k-1), \ldots, i+n(p-1)\}$ be the node set in the expanded graph corresponding to node i in the original graph. Define the additional sets, and matrices used in formulating the multimodal RCSP:

$$
\begin{aligned}
&C^k = \{(l,v) : (l,v) \text{ is an intra-mode arc in the graph for mode } k\}\\
&D^i = \{(l,v) : (l,v) \text{ is an inter-mode arc in the graph for node } i\}\\
&\eta = \bigcup_k N^k\text{: the node set for the expanded graph}\\
&\Lambda = \left(\bigcup_k C^k\right) \bigcup \left(\bigcup_i D^i\right) : \text{ the arc set for the expanded graph}\\
&G^k = (N^k, C^k)\text{: the graph for mode } k\\
&H^i = (M^i, D^i)\text{: the graph for node } i\\
&\Psi = (\eta, \Lambda)\text{: the expanded graph}
\end{aligned}
\tag{3.5}
$$

Now we define additional indices. Let $u \in \{1, ..., |\Lambda|\}$ be the index for the arc set of the expanded graph, $m^k \in \{1, ...|C^k|\}$ be the index of the intra-mode arc set for for each transportation mode k, and $z^i \in \{1, ..., |D^i|\}$ be the index of the inter-mode arc set for each node i. Also let s and t be the indices of the source and sink nodes in the original, non-expanded graph respectively. With these sets and indices we can now develop the multimodal formulation using the matrices defined below.

$$
\begin{aligned}
E^k_{l,m^k} = &\quad \begin{cases} 1 & \text{if } l \in N^k \text{ is tail of arc } m^k \in C^k \\ -1 & \text{if } l \in N^k \text{ is head of arc } m^k \in C^k \\ 0 & \text{otherwise} \end{cases} \\
F^i_{l,z^i} = &\quad \begin{cases} 1 & \text{if } l \in M^i \text{ is tail of arc } z^i \in D^i \\ -1 & \text{if } l \in M^i \text{ is head of arc } z^i \in D^i \\ 0 & \text{otherwise} \end{cases} \\
E = &\quad [E^1 \vdots E^2 \vdots \ldots \vdots E^p] \\
F = &\quad [F^1 \vdots F^2 \vdots \ldots \vdots F^n] \\
A = &\quad [E \vdots F] \\
q\_s(u) = &\quad \sum_{\substack{l \in M^s \\ A_{l,u}=1}} A_{l,u} \forall u \\
q\_t(u) = &\quad \sum_{\substack{l \in M^t \\ A_{l,u}=-1}} A_{l,u} \forall u \\
Q = &\quad \begin{bmatrix} q\_s \\ \ldots \\ q\_t \end{bmatrix}
\end{aligned} \tag{3.6}
$$

Additionally, we define the following vectors and matrices. Note that in the definitions below, the notation B' is taken to be the matrix transpose of B.

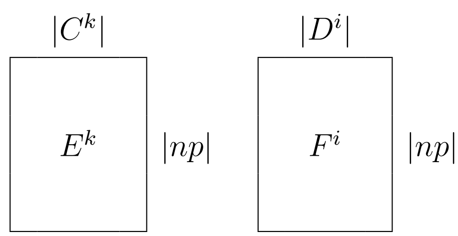
$$
\begin{array}{ll}
h: & \text{the } 1 \times p \text{ matrix of mode transportation costs in dollars per MTM} \\
r: & \text{the } 1 \times p \text{ vector of MTM available in mode } k \\
B\_s^k: & \text{the } 1 \times |C^k| \text{ vector of arc weights for mode } k \\
B\_t^i: & \text{the } 1 \times |D^i| \text{ vector of arc weights for node } i \\
B\_s = & [B\_s^1 \vdots B\_s^2 \vdots \ldots \vdots B\_s^p] \\
B\_t = & [B\_t^1 \vdots B\_t^2 \vdots \ldots \vdots B\_t^n] \\
x: & \text{the } 1 \times |\Lambda| \text{ vector of decision variables} \\
B' = & \begin{bmatrix} B\_s' \\ \ldots \\ B\_t' \end{bmatrix} \\
J_1 = & \begin{bmatrix} \frac{B\_s^1}{h(1)} & 0 & \ldots & 0 \\ 0 & \frac{B\_s^2}{h(2)} & \ldots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & & \frac{B\_s^p}{h(p)} \end{bmatrix} \\
J = [J_1 \vdots 0] &
\end{array}
\tag{3.7}
$$

Using the preceding developments we can now formulate the multimodal RCSP as:

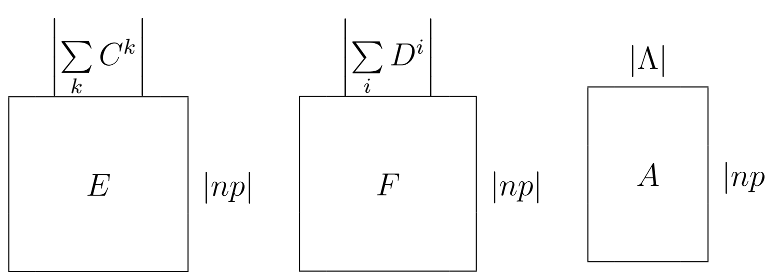
$$
\begin{array}{ll}
\text{Minimize} & Bx \\
\text{subject to:} & Ax = 0 \\
& Qx = 1 \\
& Jx \leq r \\
& x \in \mathbf{B}^{|\Lambda|}
\end{array}
\tag{3.8}
$$

3.2.2.1 Dimensional Analysis & Problem Structure. In many real-world math programming applications, the constraint matrix exhibits special structure which may be exploited to quickly solve large problems [77]. In fact, many commercial solvers perform preprocessing analysis of problems in order to determine if such structure is present in a given problem instance. These solvers will exploit the problem structure internally without intervention by the user. The constraint matrix generated in the previous section exhibits a bordered angular structure exploitable through cross decomposition which is a combination of Dantzig-Wolfe and Benders' decomposition. Before presenting the entire constraint matrix, we begin by showing consistency of dimensions among all constraint matrix subcomponents.

Each matrix, E^k, has a column for each arc contained in the intra-mode graph of mode k . Each F^i matrix has a column for every arc in the inter-mode graph on node i. Both E^k and F^k have a row for every node in the expanded graph as shown in the figures below. Note that E^k will have rows of zeros for all nodes l not in mode k $\forall k \in P$. The structure of each matrix E^k results in the matrix E having specal structure as discussed in the proceeding paragraph.



The node-arc incidence matrix, A, for the expanded graph has a row for each node and a column for each inter-modal and intra-modal arc. A comprises both E, the matrix formed by horizontally concatenating all matrices E^k and F, the matrix formed by horizontally concatenating all matrices F^i.



The following graphic shows the special block diagonal structure of the E matrix. This structure is due to the way the expanded multimodal graph is constructed by creating a “copy” of the nodes with unique labels ofr each transportation mode. Each Σ in the structural representation of E below depicts a block of non-zero elements and the 0 entries depict appropriately dimensioned blocks of 0 entries.

$$\begin{array}{cc}
\sum_k C^k & \\
\boxed{\begin{array}{ccccc} \Sigma & 0 & & 0 & 0 \\ 0 & \Sigma & & 0 & 0 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & & \Sigma & 0 \\ 0 & 0 & & 0 & \Sigma \end{array}} & |np|
\end{array}$$

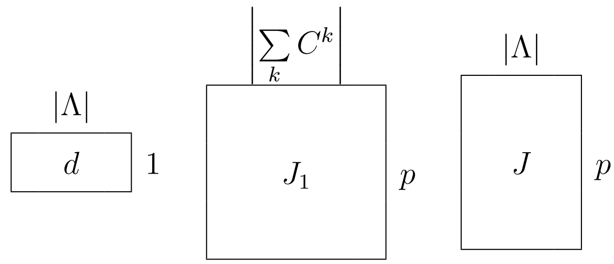
The vectors in the constraint matrix that ensure exactly one source and one sink node are chosen have a single row and a have column for every arc in the expanded graph.

$$\begin{array}{cc} |\Lambda| & \\ \boxed{q\_s} & 1 \end{array} \qquad \begin{array}{cc} |\Lambda| & \\ \boxed{q\_t} & 1 \end{array} \qquad \begin{array}{cc} |\Lambda| & \\ \boxed{\quad Q \quad} & 2 \end{array}$$

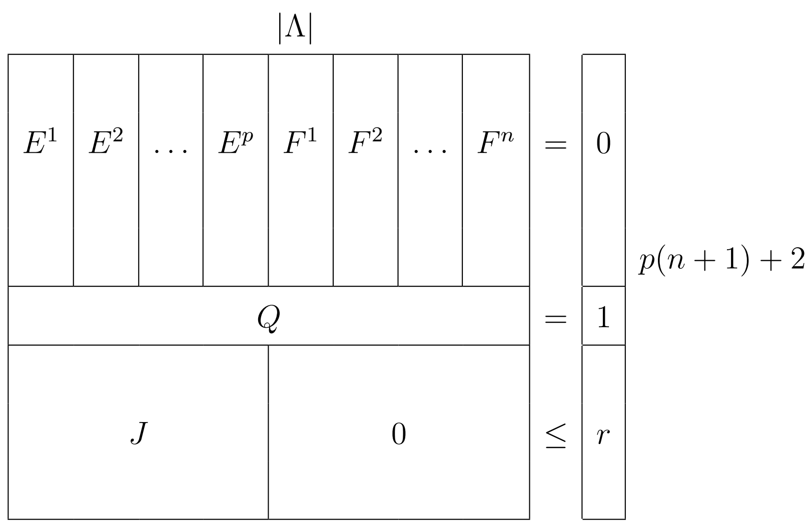
The dimensional analysis for the B_s and B_t is identical to the analysis for E and F. These vectors are provided below for completeness along with the vector,x, of decision variables.

$$\begin{array}{cc} \sum_k C^k & \\ \boxed{\quad B\_s \quad} & 1 \end{array} \qquad \begin{array}{cc} \sum_i D^i & \\ \boxed{\quad B\_t \quad} & 1 \end{array} \qquad \begin{array}{cc} |\Lambda| & \\ \boxed{\quad x \quad} & 1 \end{array}$$

The vector, d, has a column for each arc in the expanded graph. The matrix, J_1 has a row for each intra-modal arc in the expanded graph and a column for each mode of transportation appending an appropriately dimensioned matrix of zeros to J_1 yields the matrix J with a column for each arc in the expanded graph and a row for each mode of transportation.



The graphic below depicts the structure of the complete constraint matrix. The dimensions of the complete constraint matrix are $(p(n+1)+2) \times |\Lambda|$ are obtained by summing the component rows and columns.



This modeling approach represents a capacity planning at a Rough-Cut level of fidelity since it requires only operational level inputs. This method of capacity estimation is analogous to the "Capacity Planning based on Overall Factors" which uses historical data to perform mid-range to long-range planning in the context of a hierarchical production planning process. Readers unfamiliar with Hierarchical Production Planning are encouraged to reference Hax and Candea or Bitran and Hax [19, 58].

One feature of the model that is important to note is that it will always return the least cost feasible shipping option for the required capacity. This means that if a single mode option is least expensive and the fleet size for the mode is capable of providing the required number of million-ton-miles, then the solution returned will be a single mode solution.

3.3 Example Problem Formulation

In this section we explore the nature of the methodology presented in the previous section by formulating a small example problem to aid in methodology conceptualization and to show empirically, the utility of the modeling approach developed. An interesting feature of the modeling approach is that it accounts for the intermodal transfer costs. Due to this fact, the model will perform mode *selection* if a single mode of transportation is both feasible and least expensive. Alternatively, if an integrated multimodal route is less expensive, or if it is the only feasible option due to a lack of capacity in other transportation modes, then the model will select an integrated multimodal route.

The formulation presented below was constructed to demonstrate a situation where an integrated multimodal route solution is optimal. This same problem is modified slightly (the intermodal transfer weights are increased) in order to demonstrate the use of the model for single mode selection. The final characteristics demonstrated using this small example problem are the affect on capacity (and the resulting optimal solution) due to changes in capacity and planning horizon length.

3.3.1 An Example Formulation. Beginning with the single-mode graph presented in section 3.2, we expand the graph to include two additional modes of transportation. Each of the transportation modes has a different node adjacency structure. The intra-modal graphs of the two additional transportation modes are shown below. These graphs represent G^1, G^2, and G^3.

Now in order to define the complete expanded graph, the nodes must be relabeled according to $v = i + n(k-1)$ and the inter-modal graphs defined on each node

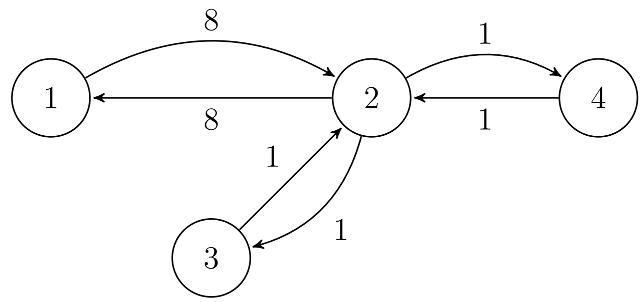


Figure 3.3: An Example Network: G^1

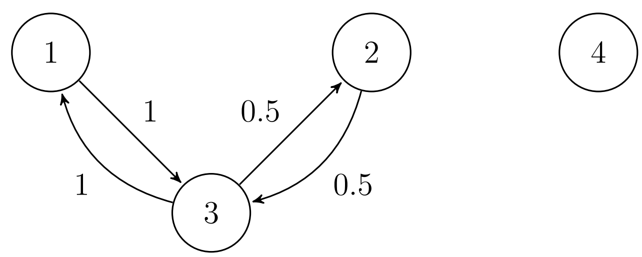


Figure 3.4: Graph for G^2

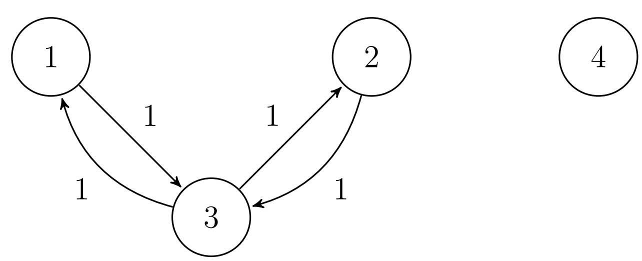


Figure 3.5: Graph for G^3

in the original graph must be generated. The inter-modal graphs are shown below. These graphs shown in Figure 3.5 (from left to right) are the graphs H^1, H^2, H^3, and H^4:

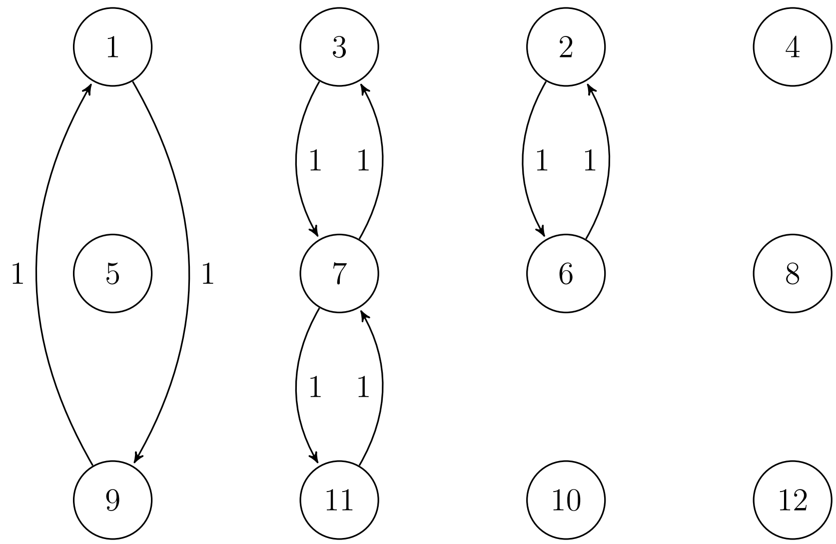


Figure 3.6: Intermodal Graphs of H^1, H^2, H^3, and H^4

By combining the intra-mode and inter-mode graphs it is now possible to specify a single graph for use in formulating the multimodal RCSP as shown in Figure 3.6.

The sets used to formulate the preceding graph are provided below for comparison and reference. In this example $n = 4$, $p = 3$.

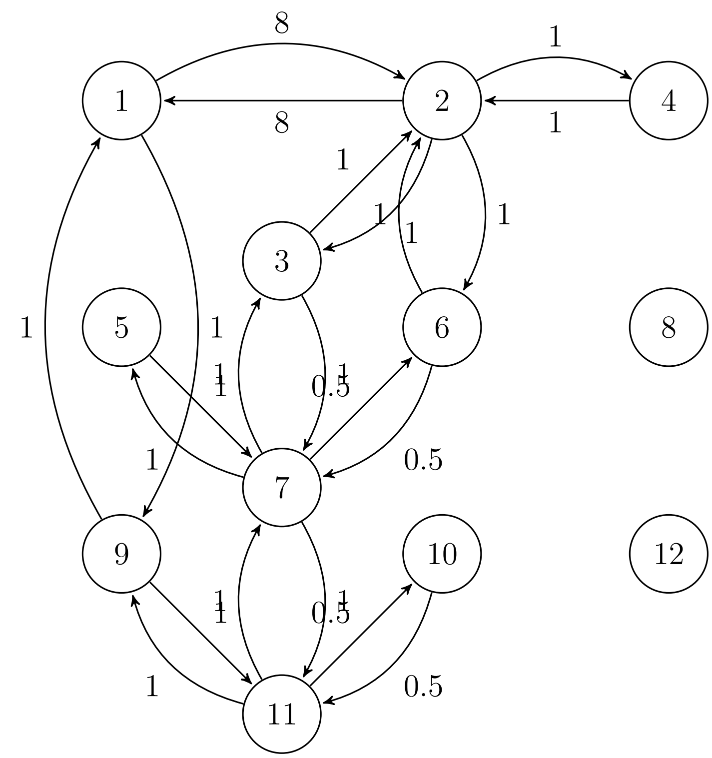


Figure 3.7: Multimodal Graph

$$
\begin{aligned}
& i \in \{1,2,3,4\}, \quad k \in \{1,2,3\} \\
& l, v \in \{1,2,3,4,5,6,7,8,9,10,11,12\} \\
& s = 1, \quad t = 4 \\
& N^1 = \{1,2,3,4\}, \quad N^2 = \{5,6,7,8\}, \quad N^3 = \{9,10,11,12\} \\
& M^1 = \{1,5,9\}, \quad M^2 = \{2,6,10\}, \quad M^3 = \{3,7,11\}, \quad M^4 = \{4,8,12\} \\
& C^1 = \{(1,2),(2,1),(2,3),(3,2),(2,4),(4,2)\} \\
& C^2 = \{(5,7),(7,5),(6,7),(7,6)\} \\
& C^3 = \{(9,11),(11,9),(10,11),(11,10)\} \\
& D^1 = \{(1,9),(9,1)\} \\
& D^2 = \{(2,6),(6,2)\}, \quad D^3 = \{(3,7),(7,3),(7,11),(11,7)\}, \quad D^4 = \{\} \\
& \eta = \{1,2,3,4,5,6,7,8,9,10,11,12\} \\
& \Lambda = \left\{ \begin{array}{c} (1,2),(2,1),(2,3),(3,2),(2,4),(4,2),(5,7),(7,5), \\ (6,7),(7,6),(9,11),(11,9),(10,11),(11,10),(1,9), \\ (9,1),(2,6),(6,2),(3,7),(7,3),(7,11),(11,7) \end{array} \right\}
\end{aligned} \tag{3.9}
$$

Indices relying upon the above set definitions are u, m^k, and z^i. These indices and the matrices they index are:

$$
E_1 = \begin{bmatrix}
1 & -1 & 0 & 0 & 0 & 0 \\
-1 & 1 & 1 & -1 & 0 & -1 \\
0 & 0 & -1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
$$

$$E_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$E_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$E = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$F_1 = \begin{bmatrix} 1 & -1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ -1 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$F_2 = \begin{bmatrix} 0 & 0 \\ 1 & -1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ -1 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$F_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$F_4 = \text{ Omitted since it is a matrix of zeros}$$

$$F = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

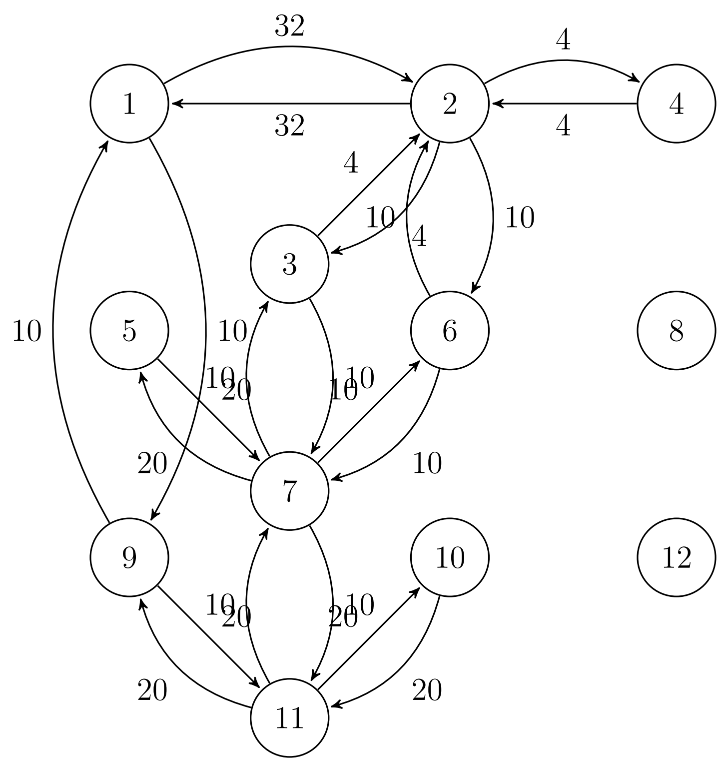
$$A = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$q\_s = \begin{bmatrix} 1\,0\,0\,0\,0\,0\,1\,0\,0\,0\,1\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0 \end{bmatrix}$$

$$q\_t = \begin{bmatrix} 0\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \end{bmatrix}$$

$$Q = \begin{bmatrix} 1\,0\,0\,0\,0\,0\,1\,0\,0\,0\,1\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \end{bmatrix}$$

Before the formulation can be completed, the distance labels applied to the multimodal network must be modified by accounting for the parameters associated with time and shipping/transfer costs. Let $h = [0.4, 2, 2]$ be the vector of shipping costs per MTM. There is an entry in the vector for each of the three transportation modes. Similarly, assume the remaining parameters are specified as follows: number of vehicles in each mode is given by the vector $[30, 21, 20]$, block speed in miles per hour is specified by $[300, 400, 500]$, UTE rate in hours per day by $[10, 12, 15]$, single trip payload for each vehicle type in units of tons is given by the vector $[20, 20, 20]$, and the productivity as an average historical percentage of trips where each vehicle type travels with a full freight load as $[.5, .5, .5]$. Additionally, let us assume we have $T = 100$ days to complete the freight movement and assume we wish to transport ten million tons of freight $\phi = 10$. Finally, assume the intermodal transshipment costs are \$1 per million tons. The values in B_t, the inter-modal arc weights, are calculated based upon the transfer cost per MTM and the number of tons (in millions) of demand (ϕ). An intermodal arc from mode 1 to node 9 ($B\_t(1)$), for example is reweighted as $B\_t(1) = h(1) * \phi = 1 * 10 = 10$. The reweighed graph represents the graph for ψ.



Modifying the arc weights in order to represent capacity in terms of MTM, we can now generate the B_s, B_t, d, J_1 and J matrices and and the r vector.

$$r = \left[\begin{array}{ccc} 90 & 96 & 150 \end{array} \right]$$

$$B\_s = \left[\begin{array}{cccccccccccccc} 32 & 32 & 4 & 4 & 4 & 4 & 20 & 20 & 10 & 10 & 20 & 20 & 20 & 20 \end{array} \right]$$

$$B\_t = \left[\begin{array}{cccccccc} 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 \end{array} \right]$$

$$d = \left[\begin{array}{cccccccccccccccccccccc} 32 & 32 & 4 & 4 & 4 & 4 & 20 & 20 & 10 & 10 & 20 & 20 & 20 & 20 & 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 \end{array} \right]$$

$$J_1 = \left[\begin{array}{cccccccccccccc} 80 & 80 & 10 & 10 & 10 & 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 10 & 10 & 5 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 10 & 10 & 10 & 10 \end{array} \right]$$

$$J = \left[\begin{array}{cccccccccccccccccccccc} 80 & 80 & 10 & 10 & 10 & 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 10 & 10 & 5 & 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 10 & 10 & 10 & 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

This completes the formulation of the constraint matrix components. It is a trivial exercise to verify consistency among the dimensions and formulate the entire constraint matrix and, thus, these final steps are omitted.

3.3.2 Mode Selection & Integration. One feature of the modeling approach important to note is that it will always return the least cost feasible shipping option for the required capacity. This means that if a single mode option is least expensive and the fleet size for the mode is capable of providing the required number of million-ton-miles, then the solution returned will be a single mode solution.

To demonstrate this capability, consider the example formulation of the previous section. The solution is given by the single modal path $1 \rightarrow 2 \rightarrow 4$. If modal transportation costs change then optimality may also change. Similarly, if resource constraints change then feasibility may change. Either of these changes has the potential to affect the optimal solution.

Now assume that the transportation cost vector changes to $h = [5, 2, 2]$. The updated values for B_t are shown below and the new optimal solution is given and interpreted.

$$B\_s = \left[\begin{array}{cccccccccccccc} 400 & 400 & 50 & 50 & 50 & 50 & 20 & 20 & 10 & 10 & 20 & 20 & 20 & 20 \end{array} \right]$$

The optimal solution now changes to the bi-modal path $5 \rightarrow 7 \rightarrow 6 \rightarrow 2 \rightarrow 4$. This path is interpreted in the physical network as beginning at node 1 in mode 2 and adding capacity to arc $(1, 3)$. At node 3 in mode 2, capacity is added to $(3, 2)$ in mode 2. Capacity is transfered from node 2 in mode 2 to node 2 in mode 1. Finally, capacity is added to arc $(2, 4)$ in mode 1.

Consider another mode price vector of $h = [5, 2, 0.5]$. This transportation price structure admits a truly multimodal optimal solution in which all three modes of transportation are used in an integrated way to provide the required capacity. The updated B_t vector for the new price structure is:

$$B\_s = \left[\begin{array}{cccccccccccccc} 400 & 400 & 50 & 50 & 50 & 50 & 20 & 20 & 10 & 10 & 5 & 5 & 5 & 5 \end{array} \right]$$

The multimodal solution corresponding to this value of h is $9 \rightarrow 11 \rightarrow 7 \rightarrow 6 \rightarrow 2 \rightarrow 4$. Interpretation of this multimodal path is similar to that for the bi-modal path and is left to the reader as a simple exercise. The next section discusses computation results of the methodology.

3.4 Results

Problems in this section are randomly generated according to several parameter inputs: Number of nodes, number of modes, density of arcs, and time and transportation asset availability. The problems are generated using MATLAB and output as Mathematical Programming System (MPS) files. The MPS file format is a standard modeling format which is recognized by major math programming solvers like LINGO and CPLEX. Models are solved using LINGO 11.0 running on a Toshiba Satellite E105 laptop computer with a 2.26 GHz Intel Core2 Duo CPU and 4 GB of RAM running 64-bit Microsoft Windows XP.

3.4.1 Parameter Selection & MPS File Generation Issues. Generating the MPS files was time consuming. In fact, MPS file generation took longer that solving the problem represented by the MPS file. In addition, file sizes for the generated files were very large. The largest file generated contained 250 nodes and 3 modes of transportation for a total of 750 nodes. The MPS file size for this problem was approximately 28 MB. The authors feel that the generated problems are of reasonable size to demonstrate the validity of the RCSP formulation.

3.4.2 Table of Results for Successfully Generated Problems. Problems were randomly generated using 90% arc density and three modes of transportation. The smallest problem has 150 total nodes (50 nodes and three transportation modes) and probelms were generated in increments of 50 nodes up to 250 ($250 * 3 = 750$ total nodes). Twenty-five instances of each problem size were generated in support of the analysis.

The results in table 3.1 include the problem MPS file size and solution time using LINGO 11.0. Twenty-five problems of each size were generated and the mean and standard deviation are provided for the time required to generate each problem size. The mean MPS file size is provided for each problem size. Standard deviation of the MPS file size adds no value to the analysis and is omitted.

Nodes	Modes	Total Nodes	MPS (MB)	Generate (sec)		Solve (sec)	
				μ	σ	μ	σ
50	3	150	1.3	14.57	0.23	0.63	0.18
100	3	300	5.02	66.48	0.82	2.50	0.45
150	3	450	11.24	201.62	2.96	7.75	4.58
200	3	600	19.87	569.56	7.87	17.08	8.13
250	3	750	30.94	1364.96	57.70	37.48	5.41

Table 3.1: Problem Generation and Solution Times

3.5 Contributions and Future Work

This article has made two original contributions to the state-of-the-art in freight capacity planning. First, we developed a new modeling approach for the rough-cut capacity planning problem by modeling this problem as a RCSP. This approach provides a time-scalable method of capacity planning which determines a least cost allocation of available capacity to a network route which meets a specified demand (in tons). Secondly, we have extended this capacity planning modeling approach into the domain of multimodal freight capacity planning. This extension allows the allocation of capacity to be either single or multimodal depending upon the cost of such an allocation. Contributions in this article demonstrate the validity of the proposed modeling method.

Future research will focus on generating larger instances of the problem to determine if exact solutions can be obtained. If exact solutions cannot be obtained in a reasonable amount of time, then a heuristic may need to be developed to solve such problems. Heuristics for solving the RCSP are categorized into three basic types: path ranking, labeling, or Lagrangian Relaxation. Various heuristics use either pure forms of these methods or a combination of two or more [32, 42, 47, 59, 65, 82, 100].

The contributions of this article to the field of operations research could be improved by additional research in the area of computer science. Specifically, how to quickly generate larger instances of the RCSP problem for solution by commercial solvers. The limitations experienced during testing for this article could be partially mitigated by applying computer science concepts to aid in the compact storage and efficient manipulation of the RCSP constraint matrix. However, the NP-complete classification of the RCSP will ultimately prevent generating problems of sufficiently large size due to the exponential problem growth as a function of the number of nodes and arcs.

One area for future research is the extension of swarm intelligence heuristics for solving the RCSP and on the identification of efficient solutions with respect to the resources of the RCSP. In the context of transportation these resources represent the consumption of time or transportation asset capacity. Identifying efficient solutions would aid decision makers in determining the best use of time and assets to meet the goals of the transportation system.

Additional research should be performed in the areas of analyzing transportation system robustness and resilience within the context of multimodal freight transportation. Multimodal integration presents a unique opportunity to reduce redundancy within a transportation mode while maintaining a passive resilience to meet demand among the other transportation modes within a system.

IV. An Ant Colony System for Solving the Resource Constrained Shortest Path Problem

4.1 Introduction

The Resource Constrained Shortest Path Problem (RCSP) is a variation on the classic Shortest Path Problem (SPP) in which the goal is to find a path of minimum total weight connecting a source node, s, and a sink node, t. The RCSP differs from the classic SPP in that it contains an additional knapsack constraint which is applied to the sum of the weights on the minimum s,t-path. Although the SPP is well solved by algorithms like Dijkstra's algorithm, Floyd-Warshall, and the Out-of-Kilter algorithm, no efficient algorithm for solving the RCSP has been developed [48].

The RCSP belongs to the class of optimization problems known as NP-Hard [54,110]. The necessity of solving the RCSP and other, equally difficult optimization problems has led researchers to develop various metaheuristics to quickly generate good solutions to these difficult problems.

The RCSP has previously been applied to rough-cut capacity planning in multimodal transportation [56] where it was noted that sufficiently large instances of the RCSP took longer to generate than to solve. The memory required to generate the constraint matrix for the RCSP imposes a restriction on the size of a problem instance that can be generated before running out of memory. Even the largest instances capable of being generated can be solved to optimality rather quickly.

In order to overcome this inherent difficulty, metaheuristics can be used to solve the problem. Typically, the amount of data required by a metaheuristic to solve a given problem instance is less than that required to generate the same problem constraint matrix for solution via a commercial solver. In this article we present an extension of the Ant System (AS) metaheuristic that incorporates new features in order to solve the RCSP.

4.1.1 Solving the RCSP. Traditional methods for solving the RCSP have been classified into three categories: Path Ranking, Labeling, and Lagrangian Relaxation [47, 82]. The method of k-shortest paths was developed by Hoffman and Pavley and later improved by Eppstein to run in $O(m + n\log n + kn)$ on acyclic graphs [42, 59].

Dynamic programming is one type of node labeling method which operates recursively. Dijkstra's Algorithm is a classic and well known example of this type of solution approach. Irnich and Villeneuve provide a labeling method for k-cycle elimination for $k \geq 3$ [65].

Lagrangian Relaxation-based methods usually begin by relaxing the resource constraints which make the RCSP NP-Hard. This relaxation is solved as a shortest path problem and then efforts are made to reduce the duality gap [32, 100].

Many heuristics to solve the RCSP have been developed based on one or more of these traditional methods. Other solution approaches involve the use of metaheuristics like Greedy Randomized Adaptive Search Procedure (GRASP) or local search-based metaheuristics like Evolutionary Algorithms, Tabu Search, Simulated Annealing, Guided Local Search, and Iterated Local Search being extended to solve the RCSP. *Constructive* metaheuristics construct feasible solutions and stop while *local search* metaheuristics iteratively improve upon incumbent feasible solutions.

4.1.2 Constructive and Local Search Metaheuristics. Evolutionary Algorithms, Tabu Search, Simulated Annealing, Guided Local Search, and Iterated Local Search are examples of metaheuristcs which are classified as either constructive or local search based. All of these metaheuristics except GRASP, which is a constructive metaheuristic, are local search based.

4.1.3 Swarm Intelligence and Ant Colony Optimization. The literature is rich with examples and applications of biologically inspired metaheuristics. Particle Swarm Optimization, Ant Colony Optimization, and Artificial Bee Colony Algorithms are newer examples of some biologically inspired optimization techniques [22].

Ant Colony Optimization is a popular and successful constructive metaheuristic which is expressly designed to solve network type problems. First developed in the 1992 dissertation of Marco Dorigo, ACO has been successfully applied to many problems capable of being formulated as shortest-path type problems [36]. The basic idea for ACO was inspired by observing the behavior of ants as they forage out of the nest for a food source. Each ant, as it travels leaves a pheromone trail which is detected by subsequent ants. Initially, ants randomly choose a direction in which to move, then subsequent ants are influenced by the pheromone trails left. The pheromones evaporate over time and therefore the pheromone trail on shorter paths connecting the nest and the food source tends to be stronger. Eventually, all ants will choose to follow the path containing the strongest pheromone scent. Empirically, it has been observed that in a majority of experiments the shortest path connecting the nest to the food source is selected by all ants after sufficient time has elapsed [50]. ACO is related to other reinforcement learning approaches through artificial pheromones and evaporation mimicking a process called stigmergy which makes pheromone trail strength available to all ants. It is assumed that the reader is familiar with the basic terminology, components, and operation of ACO algorithms. For readers requiring further background information, the 2004 book on ACO is recommended [37].

Ant Colony Optimization can include both constructive and local search elements and has been successfully extended to solve many shortest-path-type problems in the past. Problems in the areas of routing, assignment, scheduling, subset, and machine learning problems have all been addressed using some variation of ACO [14, 17, 46, 68, 99, 114]. Many other articles have been published using some variation of ACO to solve a variety of problems. An ACO survey paper by Dorigo et al. contains over one hundred references that apply ACO variations to at least eighteen different problem types in five different problem categories [34].

4.1.4 The Basic Ant Colony Optimization Metaheuristic. A Metaheuristic is a heuristic that is used to guide other heuristics in searching a solution space for an optimal solution. A heuristic is likely to become "trapped" in local optima. Hence, metaheuristics employ a variety of techniques to expand the search of the solution space and prevent premature termination of the search at local optima. However, metaheuristcs typically do not guarantee convergence to globally optimal solutions except under conditions yielding full enumeration of the solution space.

Basic ACO operates by using a group of computer agents (ants) transiting the solution space by moving from a node to adjacent nodes. The decision of which adjacent node to visit is determined through biased random selection. Ants choose randomly among remaining nodes but as the search progresses, the node transition probabilities are biased through the use of the artificial "pheromones" deposited by ants previously transiting the arc to a node. The pheromone deposits are assigned a global evaporation rate so that the amount of pheromone on an arc decreases with each passing iteration. Obviously, arcs which are visited more frequently, will have higher pheromone concentration as the search progresses. Ants select arcs with greater pheromone concentration with higher probability than those arcs with lower pheromone concentration [37]. This provides the reinforcement aspect of the search.

Each ant in the colony can leverage globally available information regarding pheromone concentrations. This indirect method of communication between ants in to colony is called "stigmergy" and is the primary mechanism of the ACO. The effects of sitgmergy are studied extensively in a 2005 report prepared for the Canadian Department of National Defense [111].

Several variations and improvements on the original ACO algorithm (referred to as simple ACO or S-ACO) have been developed. These include Ant System (AS), Elitist AS, Ant-Q, Ant Colony System (ACS), Rank-based AS, ANTS, Hyper-cube AS, and Min-Max AS (MMAS). Ant System was initially proposed by Dorigo in 1992 [33]. Ant System has inspired similar metaheuristics including Ant-Q and ACS among others. Ant-Q was developed by Gambardella and Dorigo in 1995 [44]. ACS was inspired by Ant-Q and was developed in 1996 [35,45]. It differs from Ant-Q only by the initial pheromone value applied to each arc. Empirical comparisons of these ant system algorithms indicate that the consistently best performing variants of the ant system algorithm are ACS and MMAS [89].

ACS is the most aggressive of the ACO variations and, according to Dorigo generally produces the best quality solutions for short computation times [37]. Since the method developed in this article is used as a solution generator in the context of a larger search scheme, we have selected ACS due to the empirical evidence that it produces high quality solutions in relatively short periods of time. Implementation details are discussed in the next section. We consider the RCSP problem as formulated below:

$$
\begin{array}{ll}
N = \{1, 2, ..., n\} & \text{a set of uniquely labeled nodes} \\
A = \{(i, j) : i, j \in N\} & \text{a set of directed arcs that define} \\
 & \text{adjacencies for the nodes in } N \\
G = (N, A) & \text{is a graph defined by the nodes in} \\
 & N \text{ and the arcs in } A \\
R & \text{is the quantity of resource available} \\
s & \text{is the label assigned to the source} \\
 & \text{node where supply is located} \\
t & \text{is the label assigned to the destination} \\
 & \text{node where demand is located} \\
c_{ij} & \text{is the cost of traversing arc } (i, j) \\
x_{ij} = & 1 \text{ if arc } (i, j) \text{ is included in the path} \\
 & \text{and 0 otherwise}
\end{array}
\tag{4.1}
$$

$$
\text{Minimize} \sum_{(i,j)\in A} c_{ij} x_{ij} \tag{4.2a}
$$

$$
\text{s.t.} \sum_{j:(i,j)\in A} x_{ij} - \sum_{j:(j,i)\in A} x_{ji} = b(i) \qquad \forall i \in N \tag{4.2b}
$$

$$
\sum_{(i,j)\in A} c_{ij} x_{ij} \leq R \tag{4.2c}
$$

$$
b(s) = 1, b(t) = -1, b(i) = 0 \ \forall \ i \neq s, t \tag{4.2d}
$$

In this formulation (4.2a) is the objective function, (4.2b) is the set of flow balance constraints, (4.2c) is the set of knapsack constraints and (4.2d) are the flow forcing constraints.

Merkle et al. developed an ACO metaheuristic to solve the Resource Constrained Project Scheduling Problem (RCPSP) [84]. The work is a variation of the Ant System for the Traveling Salesman Problem (AS-TSP) originally discussed by Dorigo and later by Dorigo et al. [33, 36]. Improvements to the AS-TSP made by Merkle et al. include combining two methods of pheromone evaluation, dynamic changes to the influence of the heuristic information on the probability of selection of arcs in tour construction, and an option to "forget" the best solution found so far in order to increase exploration during the search.

Hu et. al. (2010) investigated how ACO could solve navigation problems which account for multiple psychological expectations of the driver [63]. This problem is formulated as a RCSP and solved using an Ant System metaheuristic. They provide improvements to the original AS algorithm through the use of an improved transition probability rule for the ants and also an improved pheromone update scheme that consider the various driver expectations in route selection.

The remainder of this article is devoted to the development of a new Ant Colony System heuristic to solve the RCSP and to empirical testing of the heuristic's performance. An application to multimodal freight transportation capacity planning is presented and testing is executed on problem instances of various size in order to examine the relationship between parameter settings, problem size, solution time, iteration count, and solution quality. The article concludes with a discussion of the contributions made and areas of future research.

4.2 Modeling Approach

4.2.1 Selecting an Ant Colony Metaheuristic. Empirical testing suggests the best performing of the ACO variants on the Traveling Salesman Problem are ACS and MMAS [37]. Selecting one of these variants as a starting point for extension to solve the RCSP is based on comparing the observed behavior of the two variants. Dorigo and Stutzle compared the performance of ACS and MMAS on instances of the TSP contained in the TSPLIB [37]. While MMAS generally found slightly better quality solutions in long runs of the algorithm, ACS is the more aggressive search strategy and finds significantly better quality solutions for short computation times. ACS is extended and applied to the RCSP.

4.2.1.1 Solution Quality vs. Speed. Metaheuristics must be tailored to solve specific problems. In the case of ACS, there are parameters that can be tuned to control the behavior of the search. Such choices generally require tradeoff between computational performance and solution quality. Some major configuration decisions in implementation are whether or not to implement a local search strategy and if local search is used then determining pheromone update type (Darwinian vs. Lamarckian), use of data structures, and type of heuristic information about network arcs.

Implementation of a local search begins with a tour constructed by an ant (s_1) and through the local search process yields an improved path (s_2). Reinforcing the pheromone trail on s_1 is referred to as a *Darwinian* pheromone update while reinforcing the pheromone trail along the arcs corresponding to s_2 is a *Lamarckian* update.

In this research no local search is implemented following initial path construction. Eliminating local search produces a time savings of about ten percent [37]. Heuristic information is used in tour construction. The method of tour construction used in ACS determines for each ant k at city i the city j that is visited next by:

$$j = \begin{cases} \operatorname{argmax}_{l \in N_i^k} \left\{ [\tau_{il}]^{\alpha} [\eta_{il}]^{\beta} \right\}, & \text{if } q \leq q_0; \\ J, & \text{otherwise} \end{cases} \tag{4.3}$$

In 4.3, the parameter η_{ij}^k is the multiplicative inverse of the Euclidean distance between i and j, ($\eta_{ij}^k = 1/d_{ij}$). This Euclidean distance heuristic provides a more aggressive search strategy than simply using the multiplicative inverse of arc distance to provide the heuristic information. The parameter τ_{ij}^k is the strength of the pheromone on arc (i, j), N_{ij}^k is the neighborhood of ant k while at node i, q is a uniformly distributed random variable on the interval $[0, 1]$, $(0 \leq q_0 \leq 1)$, β is the tuning parameter for the weight of the heuristic value relative to the pheromone strength, and α is the pheromone weight parameter. Also let J be a random variable from the probability distribution:

$$p_{ij}^k = \frac{[\tau_{ij}]^\alpha [\eta_{ij}]^\beta}{\sum_{l \in N_i^k} [\tau_{il}]^\alpha [\eta_{il}]^\beta}, \text{ if } j \in N_i^k \tag{4.4}$$

Pheromone levels are updated after each ant has constructed a path. The two update operations applied to the pheromones are evaporation and deposit. Collectively, these two operations provide the stigmergistic aspects of the heuristic. Stigmergy is manipulated by changing the rate and method for evaporation and deposit. All ants have access to this common information and therefore communicate indirectly through updates to these stigmergic parameters. Evaporation controls how quickly the corporate "memory" of estimated arc quality is "forgotten." Deposit updates, discussed in the following section, provide a means of updating the collective corporate memory.

Finally, ASC differs form other ACO metaheuristics in three basic ways. First, it uses a more aggressive search strategy in the construction step by heavily exploiting the learned knowledge of other ants through the pheromone trails. Second, ACS only employs pheromone deposit/evaporation on the best path found so far in the search. Finally, pheromone strength is reduced on an arc immediately after it is traversed by an ant which drives exploration by making previously explored arcs less likely to be selected.

In this research, we use a designed experiment to optimize the parameter settings used in ACS as applied to the RCSP. Typical implementations of ACS assume a value of $\alpha = 1$. In our experimentation we treat α as a variable value to which optimization is applied along with the other search parameters in order to determine the best setting of all parameters relative to the RCSP.

4.2.2 Extension of ACS to Solve RCSP Problem. Ant Colony Metaheuristics are often used to construct Hamiltonian Circuits in solving the traveling salesman problem. Such a search explicitly requires that each node in the network is visited exactly one time. In constructing shortest paths it is only necessary to specify a source and sink node and perform the search until the sink node has been visited at which point the search terminates and the current path is returned as a feasible solution.

In ACS, both a global and a local pheromone update are used. The local pheromone update is done by all ants immediately after traversing an arc. The global pheromone update is done once per iteration only by the ant that has found the best path so far.

An iteration consists of all operations required for each ant in the colony to build a path, determine the best path from among these paths (s_{ib}), compare (and replace if necessary) s_{ib} to the best path so far over all iterations (s_{bs}), and complete the local and global pheromone updates, both deposits and evaporation procedures.

Global pheromone updates are applied at the end of each iteration according to the following replacement operation: $\tau_{ij} \leftarrow (i-\rho)\tau_{ij}+\rho\Delta\tau_{ij}^{bs}$, $\forall(i,j) \in P^{bs}$. In this operation, $\Delta\tau_{ij}^{bs} = 1/(C^{bs})$, where C^{bs} is the length of the best path found so far, P^{bs} is the best path so far, and $\rho \in (0,1]$ is the pheromone evaporation rate parameter. Local pheromone updates are applied according to: $\tau_{ij} \leftarrow (1-\xi)\tau_{ij} + \xi\tau_0$. In this operation $\xi \in (0,1]$ is an evaporation parameter. The local pheromone update is applied to an arc immediately after an ant crosses the arc. The effect of the local pheromone update is to decrease the desirability of an arc after it is traversed in order to drive diversification of the search and avoid search stagnation. The effect of the global pheromone update is to reinforce the pheromone on arcs that are on the best path found so far in the search. The local pheromone update prevents search stagnation by encouraging exploration of previously unvisited arcs.

Since ACS is usually applied to the TSP, the τ_0 parameter is typically set to a value of $1/(nC^{nn})$ where n is the number of cities in the TSP and C^{nn} is the length of a tour obtained using a "nearest neighbor" heuristic [37]. In shortest-path problems we do not have an *a priori* value for either n or C^{nn} and will have to estimate these parameters. Options for estimating n might include shortest path heuristics like A^* or solving the shortest path problem directly using a shortest path algorithm like Dijkstra's Algorithm. Either of these options could significantly increase computational time.

Estimates for n and C^{nn} are obtained by approximating the number of "hops" on the shortest path (n_{st}), and the length of the shortest s,t-path estimated as the Euclidean s,t distance (E^{st}). Now our estimate is taken as: $\tau_0 = 1/(n_{st}E^{st})$. n_{st} is obtained by dividing E^{st} by the average length of an arc in the network (a_l). To obtain a_l we divide the sum of all network arc distances (t) by the number of arcs in the network (m). We can calculate τ_0 directly for any network encountered as: $\tau_0 = t/(m(E^{st})^2)$.

Dorigo and Stutzle suggest using the multiplicative inverse of arc distance, $\eta_{ij} = \frac{1}{d_{ij}}$, as a heuristic measure of the desirability of arc (i, j) [37]. If we assume Euclidean distance then: $dis_{ij} = [(x_i - x_j)^2 + (y_i - y_j)^2]^{1/2}$. Another heuristic is:

$$\eta_{ij} = \begin{cases} \frac{1}{dis_{jd}} & \text{if } dis_{jd} \neq 0 \\ 1 & \text{if } dis_{jd} = 0 \end{cases} \tag{4.5}$$

Where dis_{jd} is the Euclidean distance from node j to the destination node d. This Euclidean distance heuristic for use in ACO was first proposed by Hu et. al. [63] to make the search less "myopic" at each iteration by encouraging selection of nodes that are closer to the destination node. The relative importance of the pheromone and heuristic values are mediated by the parameters α and β.

Traditional Ant Colony Optimization constructs feasible paths by accounting for node adjacency in non-complete graphs. The RCSP has the knapsack constraint that imposes an additional restriction on path feasibility. A given path is feasible only if it is also resource-feasible. Therefore, the neighborhood of the current node consists of nodes that are adjacent *and* resource-feasible.

Two mechanisms are used to ensure that resource-feasible paths are guaranteed to be constructed if they exist. First, resource consumption on partially constructed paths is maintained for each ant, making possible the identification of non-resource-feasible nodes in each neighborhood. Secondly, backtracking is implemented to allow ants to move to a previously visited node if no resource-feasible nodes exist in the current neighborhood. This approach guarantees construction of resource feasible paths if such paths exist.

4.3 Testing the ACS Metaheuristic

Our ACS metaheuristic was tested in three stages. The first stage involves determining the parameter values for the ACS. The second stage involves running ACS on problems with known optimal solutions to characterize the quality of the ACS solutions along with the speed with which they are obtained. Finally, ACS is tested on large problem instances to demonstrate that the ACS can quickly find feasible solutions to problems that cannot be generated for solution using a deterministic binary integer programming approach.

4.3.1 Selection of Parameter Values. An ACS, offers two primary objectives: "*running time*" and "*percent above the minimum objective function value.*" These response variables are affected by the problem and by the ACS parameter

Parameter Name	Description	Limits (Lower,Upper)*
α	controls the relative influence of the pheromone trail	$[0.1, 1]$
β	controls the relative influence of the heuristic information	$[1, 6]$
max iterations	the number of complete search iterations each ant colony should perform	$[10, 1000]$
num ants	the size of the ant colony used in the search	$[5, 1000]$
num neighbors	the number of nearest neighbors to consider in selecting each successive node	$[5, 25]$
q_0	pseudorandom proportional action choice rule	$[0.1, 1]$
ρ	pheromone evaporation rate	$[0.1, 0.9]$
ξ	local pheromone trail update rule	$[0.1, 1]$

* limits as suggested by Dorigo and Stutzle [37]

Table 4.1: Parameter Value Ranges

values. Design of Experiments (DOE) and Response Surface Methodology (RSM) provide tools to find input parameter values [92].

Optimizing multiple response variables can be accomplished by fitting models to each response variable and then jointly optimizing both responses using desirability functions [31]. Since ACS has stochastic components, replications are used. The focus of of our initial experimentation is to characterize the solution space. Rather than using a 2^8 full factorial design of 256 experimental runs, we opt for a central composite design (CCD) [92]. This design ensures independent estimates of all main effects and two-way interaction terms. However, some of our experimental parameters must be set to integer values and other parameters are being experimented with at the their operating extremes. Since the CCD axial points may not correspond to integer parameter values we opt to use a non-rotatable variation of the CCD called a face-centered cube design (FCD). Although the FCD is non-rotatable, it has fairly stable prediction variance throughout the design region when two center runs are used. No significant improvement in prediction variance is gained beyond two center runs and therefore we elect to use exactly two center runs in the FCD. We have selected a fractional FCD composed of a 2^{8-1} with 128 runs, 18 axial points, and two center points. The total number of experiments is 146.

Empirical testing by Hartlage and Weir indicates that the largest problem instance able to be generated for solution by a commercial solver had 750 nodes [56]. All attempts to generate larger problems failed due to memory issues. We ran the ACS experiments here using a randomly generated network with 750 total nodes and ninety percent arc density. The table below shows the fit statistics and the analysis of variance information obtained using JMP.

Not surprisingly, the model fit for for the response of "percent over optimal" is weak. This is driven mainly by the propensity of the ACS metaheuristic to converge

R^2	0.375135
$R^2_{adjusted}$	0.102916
RMSE	7.070444
Observations	146

Table 4.2: Fit Statisitcs for % over optimal

Source	**DF**	**SS**	**MS**	**F ratio**
Model	44	3031.2051	68.8910	1.3781
Error	101	5049.1084	49.9912	Prob > F
Total	145	8080.3135		0.0953

Table 4.3: ANOVA Table for % over optimal

to an optimal solution in an overwhelming number of test instances, even when varying the parameter settings. This tendency indicates that the ACS metaheuristic is fairly robust to parameter values. The plot of predicted versus actual values provides visual evidence of this assertion. Notice the appearance of a horizontal grouping of test points whose actual value was "zero." That is, these points are test points for which the predicted value was greater than the actual value. Stated another way, the metaheuristic outperformed the predicted performance in these cases. Figure 4.1 provides a concise graphical depiction of this relationship.

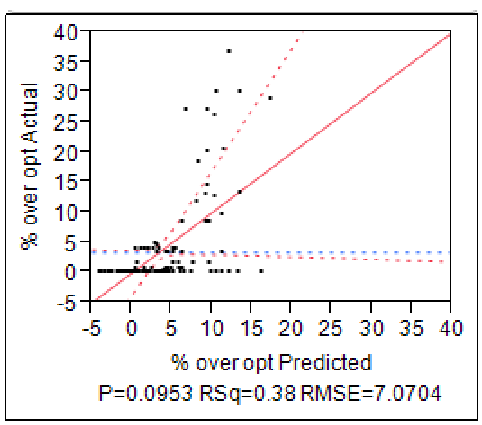


Figure 4.1: Percent over optimal for actual vs. predicted

The table below summarizes this empirical result. Notice that over sixty-percent of the time, regardless of the parameter combination used in the experiment, ACS found a solution that is within just four percent of the optimal solution.

optimal runs	49 of 146
≤ 0.01	86 of 146
≤ 0.04	88 of 146

Table 4.4: ACS Performance during Initial Experimentation

Although the relationship between predicted and actual values is statistically weak, the p-value indicates that with an α of 0.1 the model still detects variability caused by the different parameter settings.

The fit for the response of "run time" was significantly better than the model for "percent over optimal." Obviously, the run time of the heuristic is primarily a function of many controllable factors within ACS. The size of the ant colony, the number of iterations performed, and the number of neighbors in the restricted nearest neighbor list for each node are obvious drivers of this response. All values for run times are in units of seconds.

The high adjusted R^2 value and the small p-value indicate that the changes in parameter values significantly affect the variability in running time, and that the fitted model shows the strong relationship between predicted and actual values.

The parameter settings that yield maximum desirability of the two responses, are presented in the following table. These values will be used in all remaining experimentation.

R^2	0.911130
$R^2_{adjusted}$	0.872414
RMSE	42.0135
Observations	146

Table 4.5: Fit Statistics for Run Time

Source	DF	SS	MS	F ratio
Model	44	1827781.7	41540.5	23.5339
Error	101	178278.5	1765.1	Prob > F
Total	145	2006060.2		<0.0001

Table 4.6: ANOVA Table for Run Time

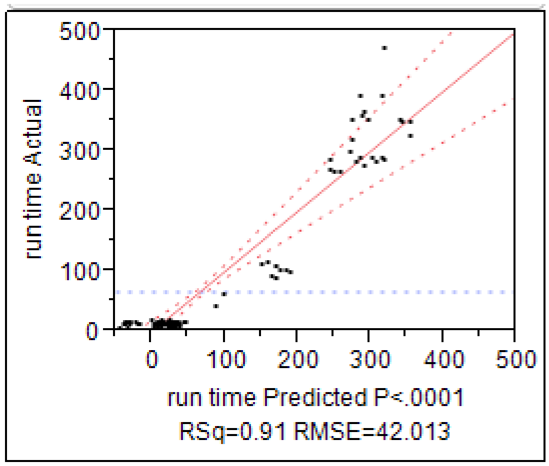


Figure 4.2: Run Time for actual vs. predicted

	Coded Value	Natural Value
α	0.1873619	0.6343
β	-1	1
MaxIterations	-0.698048	7 (rounded)
NumAnts	-1	5
NumNeighbors	-1	2
q_0	0.4758798	0.7641
ρ	1	0.9
ξ	1	1

Table 4.7: Final ACS Parameters Maximizing Desirability

The next section presents results obtained from testing the ACS metaheuristic using the parameter settings in the table above which are found to jointly maximize the desirability of the ACS metaheuristic *running time* and *solution quality*, the two responses of interest.

4.4 Results

The results in this section are based on two types of testing. The goal of the first group of tests is to provide empirical verification of the final ACS parameter settings affect on the two responses. Certain combinations of parameter settings produce significantly longer running times even on smaller problems able to be solved to optimality by solver software.

Due to the size of the constraint matrix generated for relatively small problem instances, not only was the process of problem generation time consuming, but no constraint matrix for BIP exceeding $250 * 3 = 750$ total nodes was successfully generated due to memory overrun issues. Due to this limitation, testing ACS solution quality required maintaining small problem sizes (750 total nodes) in order to obtain optimal solutions for comparison.

4.4.1 Small Example Problems. Using the settings obtained through the method of maximum desirability we'll now run ACS against LINGO to compare speed and solution quality for different problems. Density of arcs refers to the number of arcs in the network relative to the number of arcs contained in a complete graph on the same number of nodes. We use an arc density of ninety percent for purposes of experimentation. This density was selected through trial and error. Greater densities tend to produce single arc solutions which are rather uninteresting as they relate to multimodal paths while less dense solutions sometimes result in disconnected networks with no feasible solutions. In real-world problems, connectivity would be less of an issue since the input to the problem is typically a transportation network currently in use. For a complete discussion of graph completeness and density see the text by West [12]. Table 4.8 summarizes the results of this testing by providing a summary of the running time (in seconds) for ACS and BIP along with a column showing the performance of ACS in terms of solution quality (given in terms of percent over optimal).

For this experiment test we perform twenty-five replications using randomly generated networks containing 250 nodes and three modes. All networks have 90 percent arc density. The results in Table 4.8 indicate that the parameter values selected by the method of maximum desirability produce solutions are superior to those obtained in the parameter selection experiments in both running time and solution quality.

	Parameter Select Expt	Max Des Settings Expt
Num runs	146	25
Number of Optimal Runs	49 = 33.56%	12 = 48%
Runs within 1% of Opt	86 = 58.90%	14 = 56%
Runs within 3% of Opt	88 = 60.27%	20 = 80%
Worst run Pers over opt	3674.71%	11%
Mean Run Time (sec)	65.49	0.9974
StDev Run time	117.62	0.5732
99% CI for Run Time	(40.38,90.60)	(0.70,1.29)
Mean Percent over Opt	338.97%	2.00%
StDev Percent over Opt	746.50%	3.16%
99% CI for Percent over Opt	(179.58,498.36)	(0.37,3.63)

Table 4.8: Max Desirability vs. Parameter Select Expt Settings

Note that the 99% confidence intervals on the mean run time and percent over optimal for the two experiments do not overlap indicating statistical superiority of the max desirability parameter settings.

The limiting factor in the use of deterministic approaches to the RCSP is in generating the constraint matrix for the problem. The previous tests demonstrate empirically that ACS produces solutions (even removing generation time from consideration) more quickly than deterministic approaches for small problems. The remaining set of experiments is intended to demonstrate that ACS is able to quickly generate and solve problems that are too large to be generated for solution via BIP.

4.4.2 Larger Example Problems. We begin this section by solving problems with three transportation modes and 600 nodes for a total f 1800 nodes. Nodes are progressively incremented by $25 * 3 = 75$ nodes for each successive experiment. Twenty-five problems of each size with 90% arc density are solved and the results are presented in Table 4.9. These larger probelms demonstrate the applicability of this solution approach for realistically-sized problem instances.

Total Nodes	Gen (sec)	Gen $\hat{\sigma}$	Solve (sec)	Solve $\hat{\sigma}$
1800	2.2290	0.0822	3.2798	0.9335
1875	2.3810	0.0750	3.3292	0.4862
1950	2.5576	0.0705	4.4426	0.6586
2025	2.7453	0.1385	3.6931	0.4839
2100	3.0298	0.0573	4.5830	0.4793
2175	3.1771	0.1266	4.5443	0.5391
2250	3.4552	0.1202	4.5717	0.5026
2325	3.6262	0.0870	6.0689	0.7672
2400	3.8958	0.1316	5.7184	0.5939
2475	4.1083	0.1147	5.1899	0.4678
2550	4.4337	0.1224	5.1891	0.5870
2625	4.6430	0.1288	6.1895	0.4923
2700	4.9430	0.1186	8.6282	0.6248
2775	5.2524	0.1430	7.6742	0.6755
2850	5.5732	0.1367	7.0104	0.4993
2925	6.0586	0.1643	7.8940	0.6168
3000	6.1649	0.1722	7.2732	0.6885
3075	6.4917	0.1337	8.6660	0.6033
3150	6.7733	0.1803	8.0548	1.1692
3225	7.0783	0.1586	8.8562	0.6442
3300	7.4648	0.1516	9.3473	0.5789
3375	7.8142	0.2569	9.4269	0.7746
3450	8.1275	0.3497	11.1530	0.8501
3525	8.3575	0.2204	11.3913	0.9032
3600	8.8417	0.3062	10.8643	0.6984

Table 4.9: Mean and $\hat{\sigma}$ for Large Experiments

4.5 Conclusions and Future Work

This research developed an extension of the Ant Colony System metaheuristic for solving the Resource Constrained Shortest Path Problem. Empirical results show strong potential for using this metaheuristic to quickly obtain high-quality solutions to realistic sized problem instances which are too large for solution by deterministic methods.

Future research should focus on two main areas. The first is on implementing the ACS heuristic in a faster programming language like C++. Gains in speed for ACS are nearly guaranteed by implementing it in a compiled rather than an interpreted language. Secondly, further research into compact problem representation would be useful since it is problem generation due to computer memory overruns rather than solution speed that limits the problem size able to be solved.

V. *Finding Near Efficient Solutions to Lexicographic, Multi-Objective, Resource Constrained Shortest Path Problems*

5.1 Introduction

Capacity planning in transportation is an integral and important function for mid-range planning or fleet utilization. Rough-Cut Capacity Planning (RCCP) in multimodal freight transportation has previously been addressed by Hartlage and Weir and is solved using either deterministic binary integer programming solvers or metaheuristics [56]. Since the RCSP is known to be NP-Hard, sufficiently large problem instances require fast running metaheuristics to quickly provide high-quality solutions. A new metaheuristic, ACS-RCSP, based on the Ant Colony System (ACS) traditionally used to solve the Traveling Salesman Problem (TSP) was developed by Hartlage and Weir [55].

In RCCP for multimodal freight, three controllable parameters ultimately determine resource requirements and constraint right-hand-sides (RHS). These parameters are:

- number of tons (in millions) of freight to transport,
- number of transportation assets available in each transportation mode, and
- length (in days) of the planning horizon.

We use the term "controllable" to distinguish between parameters that directly affect the RHS (or arc weights) and are completely determined by the decision maker, and those parameters that affect the RHS but are primarily determined by factors not directly under the control of the decision maker. Parameters in the latter category are referred to as "uncontrollable." Some examples of parameters in this category include:

- productivity of each transportation asset type,

- utilization rate of each transportation asset type, and
- block speed of each transportation asset type.

To illustrate these parameters, consider asset utilization or "UTE rate." The UTE rate for aircraft is defined by Air Mobility Command's Air Mobility School as "the total hours of capability a fleet of airlift aircraft can produce in a day expressed in terms of per primary authorized aircraft." Similar definitions are available for other modes of transportation.

Note that higher UTE rates are not necessarily "better" in any sense of the word. A decision maker may elect to increase UTE to an unsustainable "surge" UTE rate for a short period of time but this is a strictly short-term situation.

UTE rate is affected by many factors. A list of twenty four different factors affecting aircraft UTE rate illustrates that some of the factors influencing UTE rate are controllable and some are not. This list is provided for ease of reference and in order to clearly illustrate the point that UTE rate is not directly controllable:

Average Ground Time	Average Mission Time and Leg Length
Airspeed En Route	Crew Ratio
Crew Availability	Crew Augmentation policies
Crew Stage Base policies	Active and Reserve Force Mix
Reserve Call-Up Schedule	Spares ande Resupply Availability
Maintenance Manpower	Scenario Resource Constraints
Ramp Space Constraints	PAA Airframes
JCS Withhold Levels	Aerial Refueling Policies
Air Traffic Control Delays	Political and diplomatic clearance delays
Weather	Airfield operating hours
Mission Capable Rate	Aircraft Reliability
Aircraft Maintainability	Aircraft Availability
Aircraft Generation Schedule	Sch. Maint. Interval and Duration

Table 5.1: Table of UTE Factors

5.1.1 Notions of Optimality. Multicriteria optimization problems define optimality in several ways. We must first define how the various objective components are to be compared before a notion of optimality or efficiency can be applied.

Objectives are compared explicitly by ordering them according to some set of criteria. For a complete discussion on ordering objectives and types of orders, refer to Ehrgott [40].

We are concerned with different lexicographic orderings of the objective components. Lexicographic ordering assigns a priority to each objective component so that a small increase in a higher priority component outweighs a large increase in a lower priority component. Examining the efficient solutions that result from different objective component prioritizations is a major item if interest in this research.

Depending upon the size of the RCSP problem instance (number of network nodes and density of network arcs), it may or may not be practical to locate these efficient solutions using a binary integer programming (BIP) approach. In cases where the problem instance is too large for BIP to be practical a metaheuristic, such as the ACS-RCSP may be used to generate near-optimal solutions [55]. In cases where a suboptimal solution is returned by ACS-RCSP, we consider such solutions to be "near-efficient".

5.1.2 Problem Statement. The problem under investigation in this research is the location and identification of efficient solutions to the RCSP for multimodal freight transportation. Locating such solutions requires defining some notion of optimality. In the event that a lexicographic ordering of the transportation asset types is provided, near-efficient solutions are generated with respect to the ordering. If no objective ordering is provided, the near-efficient solutions are generated with respect to minimizing costs by considering the price, in dollars, per million-ton-mile for each transportation mode, and as such an implicit lexicographic objective ordering is used.

5.1.3 Review of the Literature.

5.1.3.1 Multimodal Freight Capacity Planning. Capacity planning in multimodal freight transportation has received relatively little attention in the literature to date [79]. In fact, much of the previous research addressing multiple transportation modes focuses on mode selection rather than integration. A recent survey of the transportation mode choice and carrier selection literature reviewed over ninety separate references spanning a period of over forty years [83]!

The 2005 article by Park and Regan developed a bi-level programming formulation for multimodal freight capacity planning. However the approach uses an "average load factor" for each transportation mode and is therefore not additive with respect to capacity over the entire fleet [96].

In 2006 Sun et al. use an existing linear programming model called MAX-CAP which seeks to maximize system capacity for multiple origin-destination pairs subject to resource and capacity constraints. They analyze system "flexibility" and performance under degradation but consider only a single mode of transportation, movement by rail [104].

Morlok and Riddle address the maximization of multimodal system capacity by defining "traffic lanes" that represent aggregated paths and modes. Aggregation is done based on origin/destination pairs that are geographically close to each other. Capacity of the system is estimated by summing over the aggregated traffic lanes [88].

Unnikrishnan and Waller analyze the effect of of demand and capacity uncertainty on rail network capacity, seeking to maximize the fraction of demand satisfied for each origin/destination pair in the network [106].

Cambridge Systematics Inc. published a manual that discusses multimodal networks for both freight and passenger transportation. This manual uses multimodal "corridors" which are essentially channel routes that utilize multiple transportation modes. Capacity estimation of the network is done by analyzing the corridor capacity by analysis of individual link capacity in each multimodal corridor [105].

Andersen and Crainic address the design and optimization of a multimodal scheduled service network. The approach used in this article is to maximize system throughput by coordinating and optimizing the departure times for multiple vehicle fleets utilizing a common network [4]. No literature source reviewed addressed identification of efficient solutions under different parameter value assumptions.

5.1.3.2 RCSP Classification and Solution. The basic RCSP, formulated below, is NP-Complete as shown by Handler and Zang [54]. To simplify the presentation of the single mode resource constrained shortest path formulation the following parameters and vectors are defined: $N = \{1, 2, ..., n\}$ is the set of nodes representing shipping locations, $C = \{(l, v) : (l, v) \text{ is an arc in the graph}\}$ is the set of feasible directed arcs defined over the elements in N, $G = (N, C)$ is the graph consisting of nodes in N and edges in C, and E is the $n \times |C|$ node-arc adjacency matrix where each column of E has a 1 in the l^{th} row and a -1 in the v^{th} row corresponding to adjacency matrix element E_{lv}. Now define E_{lv} along with the following sets, matrices, and scalars as:

$$E_{lj} = \begin{cases} 1 & \text{if } l \in N \text{ is the tail of arc } j \in C \\ -1 & \text{if } l \in N \text{ is the head of arc } j \in C \\ 0 & \text{otherwise} \end{cases} \quad \forall \quad l \in N,\, j \in C$$

$$q\_s(j) = \begin{cases} 1 & \text{if } E_{sj} = 1 \\ 0 & \text{otherwise} \end{cases} \quad \forall \quad j \in C$$

$$q\_t(j) = \begin{cases} 1 & \text{if } E_{tj} = -1 \\ 0 & \text{otherwise} \end{cases} \quad \forall \quad j \in C$$

$$Q = \begin{bmatrix} q\_s \\ \dots \\ q\_t \end{bmatrix} \tag{5.1}$$

$h :$ the cost (in dollars) per MTM

$r :$ the number of MTM available

$B\_s :$ the $1 \times |C|$ vector of arc weights

$x :$ the $1 \times |C|$ vector of decision variables

$$J = \frac{B\_s}{h}$$

The single mode formulation can now be expressed as:

$$
\begin{aligned}
\text{Minimize} B\_s\ x & \qquad (5.2a)\\
\text{subject to:} Ex = 0 & \qquad (5.2b)\\
Qx = 1 & \qquad (5.2c)\\
Jx \leq r & \qquad (5.2d)\\
x \in \mathbf{B}^{|C|} & \qquad (5.2e)
\end{aligned}
$$

In this formulation 5.2a is the objective function, 5.2b are the transshipment portion of the flow balance constraints, 5.2c are the source and sink portion of the flow balance constraints, 5.2d are the knapsack constraints, and 5.2e are the flow forcing constraints.

The RCSP is a variation on the classic Shortest Path Problem (SPP) in which the goal is to find a path of minimum total weight connecting a source node, s, and a sink node, t. The RCSP adds an additional knapsack constraint which is applied to the sum of the weights on the minimum s,t-path. Although the SPP is well solved by algorithms like Dijkstra's algorithm, Floyd-Warshall, and the Out-of-Kilter algorithm, no efficient algorithm for solving the RCSP has yet been developed [48].

Formulating the RCCP in multimodal freight networks as an instance of the RCSP was first accomplished by Hartlage and Weir in 2012 [56]. The original formulation is shown below for reference.

We first define the following sets and parameters to generalize the single mode for addressing multimodal freight transportation. Let $P = \{1, 2, \ldots, p\}$ be the set of available transportation modes and $k \in P$ be be the index for these modes. Also $N = \{1, 2, \ldots, n\}$ be the set of nodes representing nodes, as in the single-mode formulation with $i \in N$ being the index for the nodes. Expanding the original graph as described in the previous paragraph results in a node set with np nodes.

The indices for the node set of the expanded graph are $l, v \in \{1, \ldots, np\}$. Let $N^k = \{1 + n(k-1), \ldots, n + n(k-1)\}$ be the node set for mode k in the expanded graph. Finally, let $M^i = \{i, \ldots, i + n(k-1), \ldots, i + n(p-1)\}$ be the node set in the expanded graph corresponding to node i in the original graph. Define the additional sets, and matrices used in formulating the multimodal RCSP:

$$
\begin{aligned}
& C^k = \{(l,v) : (l,v) \text{ is an intra-mode arc in the graph for mode } k\} \\
& D^i = \{(l,v) : (l,v) \text{ is an inter-mode arc in the graph for node } i\} \\
& \eta = \bigcup_k N^k\text{: the node set for the expanded graph} \\
& \Lambda = \left(\bigcup_k C^k\right) \bigcup \left(\bigcup_i D^i\right)\text{: the arc set for the expanded graph} \\
& G^k = (N^k, C^k)\text{: the graph for mode } k \\
& H^i = (M^i, D^i)\text{: the graph for node } i \\
& \Psi = (\eta, \Lambda)\text{: the expanded graph}
\end{aligned}
\tag{5.3}
$$

Now we define additional indices. Let $u \in \{1, ..., |\Lambda|\}$ be the index for the arc set of the expanded graph, $m^k \in \{1, ...|C^k|\}$ be the index of the intra-mode arc set for for each transportation mode k, and $z^i \in \{1, ..., |D^i|\}$ be the index of the inter-mode arc set for each node i. Also let s and t be the indices of the source and sink nodes in the original, non-expanded graph respectively. With these sets and indices we can now develop the multimodal formulation using the matrices defined below.

$$
\begin{aligned}
E^k_{l,m^k} = & \begin{cases} 1 & \text{if } l \in N^k \text{ is tail of arc } m^k \in C^k \\ -1 & \text{if } l \in N^k \text{ is head of arc } m^k \in C^k \\ 0 & \text{otherwise} \end{cases} \\
F^i_{l,z^i} = & \begin{cases} 1 & \text{if } l \in M^i \text{ is tail of arc } z^i \in D^i \\ -1 & \text{if } l \in M^i \text{ is head of arc } z^i \in D^i \\ 0 & \text{otherwise} \end{cases} \\
E = & [E^1 \vdots E^2 \vdots \ldots \vdots E^p] \\
F = & [F^1 \vdots F^2 \vdots \ldots \vdots F^n] \\
A = & [E \vdots F] \\
q\_s(u) = & \sum_{\substack{l \in M^s \\ A_{l,u}=1}} A_{l,u} \forall u \\
q\_t(u) = & \sum_{\substack{l \in M^t \\ A_{l,u}=-1}} A_{l,u} \forall u \\
Q = & \begin{bmatrix} q\_s \\ \ldots \\ q\_t \end{bmatrix}
\end{aligned}
\tag{5.4}
$$

Additionally, we define the following vectors and matrices. Note that in the definitions below, the notation d' is taken to represent the matrix transpose of d.

$h :$ the $1 \times p$ matrix of mode transportation costs in dollars per MTM

$r :$ the $1 \times p$ vector of MTM available in mode k

$B\_s^k :$ the $1 \times |C^k|$ vector of arc weights for mode k

$B\_t^i :$ the $1 \times |D^i|$ vector of arc weights for node i

$$B\_s = [B\_s^1 \vdots B\_s^2 \vdots \ldots \vdots B\_s^p]$$

$$B\_t = [B\_t^1 \vdots B\_t^2 \vdots \ldots \vdots B\_t^n]$$

$x :$ the $1 \times |\Lambda|$ vector of decision variables

$$B' = \begin{bmatrix} B\_s' \\ \ldots \\ B\_t' \end{bmatrix}$$

$$J_1 = \begin{bmatrix} \frac{B\_s^1}{h(1)} & 0 & \ldots & 0 \\ 0 & \frac{B\_s^2}{h(2)} & \ldots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & & \frac{B\_s^p}{h(p)} \end{bmatrix}$$

$$J = [J_1 \vdots 0] \tag{5.5}$$

Using the preceding developments we formulate the multimodal RCSP as:

$$\begin{aligned} \text{Minimize} \quad & Bx \\ \text{subject to:} \quad & Ax = 0 \\ & Qx = 1 \\ & Jx \leq r \\ & x \in \mathbf{B}^{|\Lambda|} \end{aligned} \tag{5.6}$$

This multimodal formulation of the RCSP is the focus of the methodology that follows. Similar research efforts have used metaheuristics to solve a similar problem called the resource constrained project scheduling problem (RCPSP). The RCPSP is similar to the RCSP in that it includes constraints on total available resources but differs in that its objective is to minimize the total time required to complete a series of tasks (called the "makespan") and also includes precedence constraints on the tasks.

Calhoun used a modified version of Van Hove's original Multi-modal resource constrained project scheduling problem with generalized precedence constraints (MMG-PRCPSP) to model the assignment of aircraft to targets. He formulated the MMG-PRCPSP as a lexicographic ordered goal program to minimize the number of targets not covered. He solved the problem using Tabu Search [25, 62]. In the context of Calhoun and Van Hove's research "multi-modal" refers to the number of different ways or "modes" in which a particular task can be accomplished. This is an important distinction since the term "multi-modal" is used in this article to mean the movement of freight using two or more modes of transportation.

Coello Coello et al. develop a method for incorporating multiple objectives into particle swarm optimization (PSO) by generating uniformly distributed points along a Pareto front [27].

Nasiri develops a pseudo PSO metaheuristic for solving the RCPSP [94]. The metaheuristic uses path relinking to allow particles to move toward optima (local and global).

Merkle et al. developed an Ant Colony Optimization metaheruistic for the RCPSP [84].

5.2 Methodology

The methodology developed here is intended to support tradeoff analysis in capacity planning decisions as they relate to fleet mix, total cost, and time. In particular, the first-of-kind methodology developed here allows for variable time resolution modeling (selection of time interval length and units), and provides insight into the value, in dollars, of additional time or transportation assets. Typical questions that may be addressed by the methodology developed here are:

- What is the minimum number of days, weeks, or months in which to ship x tons of freight with the specified mix of assets?
- What is the lexicographic minimum asset mix required to ship x tons in y days, weeks, or months?
- What is the minimum cost asset mix to ship in a given number of days, weeks, or months?
- Is the current asset mix minimal? That is, are there excess assets in the current asset mix that could be re-purposed for another mission?

Although there are certainly other questions that can be addressed using this new methodology, these typify questions a decision maker might be interested in regardless of whether the application is to military or civilian transportation.

5.2.1 Finding Near-Efficient Solutions. In theory the Pareto front is continuous and therefore contains an infinite number of efficient solutions. In some applications, such as the application to multimodal transportation discussed in this paper, the Pareto front is discrete. However, the number of discrete Pareto points on the frontier can be quite large. Consider a relatively small number of transportation modes and number of assets in each mode. Given 7 transportation modes and 20 of each transportation asset type, the number of potential efficient solutions is $20^7 = 1,280,000,000$. Finding, and comparing all of these points is time consuming and for large enough problems quickly becomes impractical. Typically, there are three items to consider when evaluating multi objective optimization problems [118].

1. Minimization of the distance between the near-Pareto points found and the global Pareto front.
2. Maximize the spread of the near-efficient solutions found to provide a smooth and uniform near-Pareto front.
3. Maximize the number of elements of the global Pareto front found.

We use an ACS metaheuristic developed by Hartlage and Weir to aid in generating solutions for this article [55]. The results included an optimization of the ACS input parameters to jointly optimize both the running time and the solution quality. These previous results suffice for addressing the first and third item in the list above. In this article, our focus is on finding uniformly distributed, near-efficient solutions.

5.2.2 Distance between Solutions and Coverage of the Pareto Front. By way of an example, consider route choices generated by an internet-based route mapping engine. Initially, a start and an end point are specified. The engine provides a driving route between the specified origin and destination. However, alternative routes are often specified. In order to ensure that an alternate route differs from the original route by more than simply cutting through a parking lot, the engine must employ a measure of distance between successive solutions.

This article considers a two dimensional objective space. The objectives considered are time (in days) and cost (in dollars). For a fixed quantity of freight, there is an inherent tradeoff between delivery time and cost. Increasing time reduces the number of transportation assets and the associated cost of capacity. Decreasing the time requires the use of more transportation assets and is likely to involve faster and more expensive transportation modes, measured in dollars per million-ton-mile, to meet shorter timelines. Shipping freight by air is an example of a relatively fast but expensive mode of transportation. Increasing time will decrease the number of transportation assets required to provide required capacity and will tend to require less expensive modes of transportation.

We can measure how uniformly the solutions are distributed in the objective space by measuring the range variance of the near-efficient solutions in the objective space. The metric for measuring range distance proposed by Schott is [103]:

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\overline{d} - d_i\right)^2} \tag{5.7}$$

Where i and j are indices of the total number of efficient solutions generated, n, and d_i is the distance from a solution, i, to the nearest solution j for $j = 1, ..., n$. Mathematically, d_i is expressed as:

$$d_i = min_j \left(\left|f_1^i(\overrightarrow{x}) - f_1^j(\overrightarrow{x})\right| + \left|f_2^i(\overrightarrow{x}) - f_2^j(\overrightarrow{x})\right|\right) \forall i, j = 1, 2, ..., n \tag{5.8}$$

Also, let $\overline{d}$ be the arithmetic average of all d_i and f^k be vectors for $k = 1, 2, ..., n$. This metric allows us to evaluate the uniformity of near-efficient solutions. Coverage of the Pareto front is another issue which requires discussion. Coverage can be determined by selecting two values in the objective space. For a fixed quantity of freight, selecting minimum and maximum values for time determines the size and location of the Pareto front segment of interest. Setting a minimum value for time (T_{min}), we obtain the highest cost near-efficient solution (C_{max}). Setting a maximum allowable time (T_{max}), we obtain the lowest cost near-efficient solution (C_{min}).

5.2.3 Generalized Range Variance for Lexicographic Objective Ordering. The definition of range variance given by Schott is only sufficient to measure the range variance in a two dimensional objective space. In the case of lexicographic ordering of m components we require range variance measured in all m dimensions. Generalizing range distance requires us to redefine the quantity d_i as:

$$d_i = min_j \left(\sum_{k=1}^{m}\left|f_k^i(\overrightarrow{x}) - f_k^j(\overrightarrow{x})\right|\right) \forall i, j = 1, 2, ..., n \tag{5.9}$$

Defining d_i in this way we can now provide range variance not only in instances where the objective space has the dimensions of cost and time but also in situation where we are interested in an objective space having dimensions of time and lexicographic order of all transportation asset types.

5.2.4 The Value of Dominated Solutions. In the context of capacity planning for multimodal transportation, global Pareto solutions have no excess capacity. Stated in another way, the loss of a single asset from among the fleet proposed in the Pareto solution causes the problem to be come infeasible for the current value of T. In some applications and situations, a decision maker may wish to include reserve capacity in the planning phase in order to account for reliability issues with the fleet or late departures of assets due to weather delays for example. Making a decision to operate with excess capacity is equivalent to selecting a dominated point (non-efficient solutions) in the objective space. Therefore, presentation of non-efficient solutions is beneficial to decision makers desiring this additional capacity.

5.2.5 Methodology Development. The input required for this methodology include parameters and data which describe the transportation mode characteristics and the configuration of the underlying network. Also, the maximum number of each transportation asset type (mode) is a required input. An overview of the process is provided here for reference.

- Provide required parameter and data inputs.
- Determine the minimum number of days required to ship the required tons of freight with the provided asset mix.
- Using data provided for number and length of time intervals, generate near-efficient solutions for each of the interval endpoints.
- Present the full set of near-efficient solutions to the decision maker.

5.2.5.1 *Parameters and Data.* There are a number of inputs that are required to initialize the problem. The following parameters determine characteristics of the network.

- h_k: The transportation cost, in dollars, per million-ton-mile for mode k.
- $LexObjOrder$: The lexicographic objective ordering for all modes of transportation $[1 : m]$.
- $UTErate$: The fleet utilization rate, in hours, for each transportation mode.
- $Productivity$: The fleet productivity, or percentage of the time that assets do not move while empty (for repositioning) for each transportation mode.
- $BlockSpeed$: The average speed, door-to-door, for each transportation mode.
- $Payload$: The single trip, payload of each transportation mode in tons.
- $XFerCost$: The cost, per million tons, to transfer freight between two modes of transportation (could be unique for each mode pairing).
- s: The source node or point of embarkation (POE) for the system.
- t: The sink node or point of debarkation (POD) for the system.
- rc: The required system capacity in millions of tons.
- $IntSize$: The number of days in a time interval.
- $IntNum$: The number of time intervals to analyze.
- $MaxAssets$: A vector, $[a_1, a_2, ..., a_m]$ of the number of transportation assets available in each mode of transportation.
- T: A scalar denoting the number of days allowed to ship the amount of freight specified by rc.

The next set of parameters determine the search characteristics for the ACS-RCSP metaheuristic:

- $NumNeighbors$: The number of nearest neighbors on the reduced nearest neighbor (NN) list. A reduced NN list reduces computation time by as much as ten percent [37].
- α: pheromone weight factor on the interval $[0, 1]$ reduces influence of the pheromones.
- β: heuristic weight factor in the interval $[0, 1]$, a lower β reduces the influence of the heuristic information.
- T_0 weight parameter for pheromone update operation
- q_0: An intensification/diversification parameter in the interval $[0, 1]$, a higher q_0 means more exploitation of learned knowledge while a lower q_0 means more exploration of arcs.

5.2.5.2 Parameters and Transportation Modes. We can conveniently represent the time and transportation asset input variables for each problem instance as a single vector. By convention, the first variable in the vector is the parameter T, in days. The remaining values in the vector indicate how many of each transportation mode asset type is available. These variables are captured in $MaxAssets$ defined in the previous section. For the purposes of this methodology, a transportation mode is a grouping of homogeneous transportation assets with access to a common network structure (nodes with adjacencies defined). Defining modes in this way allows greater flexibility in accurately describing the transportation system. For instance, although it might be tempting to classify all air transportation assets as belonging to a single mode of transportation, such a classification does not accurately capture the differences between the asset types

5.2.5.3 Determining Minimum Days Required. Determining the minimum number of days required to provide necessary system capacity involves producing an infeasible solution and using bisection to measure the infeasibility in terms of days. Producing an infeasible solution is relatively easy. However, we are interested in generating an infeasible solution using the maximum number of transportation assets available in all modes of transportation. To do this, we set $T = 1$ and solve the problem. From this infeasible solution, we implement a reverse and forward bisection routine to determine the minimum number of days, T, required to obtain a feasible solution with $MaxAssets$. First, we must define a function which determines the feasibility of a RCSP provided inputs of T and $MaxAssets$ as follows:

$$f(Time, AssetMix) = \begin{cases} 1 & \text{if RCSP admits a feasible solution} \\ -1 & \text{otherwise} \end{cases} \tag{5.10}$$

The function f used in this methodology is ACS-RCSP or a binary integer programming solver. ACS-RCSP is the Ant Colony System metaheuristic for solving the RCSP originally developed by Hartlage and Weir [55]. The reverse and forward bisection routine pseudo-code is described in Figure 1:

This algorithm, $FBintBisect$, is similar to the bisection method used to find roots of a function [24], however, it has been modified to include two additional features. First, it includes a preprocessing step that "reverse" bisects, increasing the value of T until a feasible solution is returned. Second, it uses the function $ceiling(x)$ which rounds fractional values, x, to the nearest integer value in the direction of positive infinity.

5.2.5.4 Generating Near-Efficient Solutions. The value returned by the bisection algorithm, $minT$, is the minimum number of days required in order for the problem to be feasible with the assets available in $MaxAssets$. Now, beginning from $minT$ we generate the number of near-efficient solutions specified by $IntSize$ and $IntNum$. The subroutine to generate these near-efficient solutions requires first developing three other utility subroutines: "$StValInc$", "$AssetMix$",

Algorithm 1 Forward-Backward Integer Bisection Subroutine

$L \leftarrow 0$
$H \leftarrow T$
$x \leftarrow f(H, Assets)$
while $x = -1$ **do**
  $H \leftarrow 2 * H$
  $x \leftarrow f(H, Assets)$
end while
$M \leftarrow ceiling((L + H)/2)$
while $H - L \geq 2$ **do**
  if $x = 1$ **then**
    $H \leftarrow M$
    $M \leftarrow ceiling((L + H)/2)$
  else if $x = -1$ **then**
    $L \leftarrow M$
    $M \leftarrow ceiling((L + H)/2)$
  end if
  $x \leftarrow f(H, MaxAssets)$
  $T \leftarrow M$
end while
return T

and "$FeasAssetMin$". $StValInc$ takes an infeasible solution and determines by how much each of the transportation asset types and time would have to be increased independent of one another in order to achieve feasibility. In essence, $StValInc$ provides a measure of by how much the problem is infeasible with respect to each of the controllable inputs. $AssetMix$ increases system capacity by adding assets (up to that specified in $MaxAssets$) in non-decreasing order with respect to the lexicographic ordering (or by cost per million ton mile if no objective ordering is provided) until a feasible solution is returned. Once this feasible solution is found, $FeasAssetMin$ determines if the asset mix returned by $AssetMix$ contains any excess assets. Any excess assets are removed from the asset mix so that the final asset mix returned by $FeasAssetMin$ is *minimal*, in that it is both necessary and sufficient for feasibility at the current value of T. $StValInc$ is described in Algorithm 2, $FeasAssetMin$ in Algorithm 4, and $AssetMix$ in Algorithm 3.

Algorithm 2 Subroutine for Determining Initial Increase to Attain Feasibility

$StValInc \leftarrow zeros(1, length(MinAssets))$
for $i = 1 : m$ **do**
  $StValInc(i) \leftarrow FBintBisect(Time, MinAssets)$
end for
return $StValInc$

For convenience in expressing $AssetMix$ in pseudocode define the following variables:

- $R(x)$: The number of additional assets in mode x required to make the problem feasible.
- $A(x)$: The number of additional assets in mode x available.
- $U(x)$: The number of assets in mode x currently used (Note: total assets in mode $x = A(x) + U(x)$).

Now, let $MinAssets$ be an m-vector of ones. Notice that the choice of $IntSize$ and $IntNum$ determine the distance between efficient solutions and the number of solutions generated, respectively. When combined, the selection of these parameters determines both fidelity and coverage of the decision space. The range variance provides a measure of the uniformity of distribution of the points generated in the objective space. The routine for generating efficient solutions is described in Algorithm 5.

5.3 Example Problems

We demonstrate the method of finding near-efficient solutions in this section. Results are presented as a continuum from the most costly, shortest time required solution to the least costly most time required solution. For each solution, the path, cost in dollars, and the asset mix are provided.

5.3.1 Randomly Generated Examples. The test network consists of 50 nodes, three modes of transportation, and the node positions are randomly generated

Algorithm 3 Subroutine for Determining the Asset Mix

```
if lexOrd = 1 then
    while "no feasible solution is returned" do
        x = "index of least important transportation mode such that" A(x) > 0
and R(x) > 0
        if R(x) <= A(x) and A(x) > 0 then
            U(x) ← U(x) + R(x)
            A(x) ← A(x) − R(x)
        else if R(x) > A(x) and A(x) > 0 then
            U(x) ← U(x) + A(x)
            A(x) ← 0
        end if
        for i = 1 : m do
            R(i) = StValInc(Time, U(i))
        end for
    end while
else if lexOrd = 0 then
    while "no feasible solution is returned" do
        x = "index of least expensive transportation mode such that" A(x) > 0
and R(x) > 0
        if R(x) <= A(x) and A(x) > 0 then
            U(x) ← U(x) + R(x)
            A(x) ← A(x) − R(x)
        else if R(x) > A(x) and A(x) > 0 then
            U(x) ← U(x) + A(x)
            A(x) ← 0
        end if
        for i = 1 : m do
            R(i) = StValInc(Time, U(i))
        end for
    end while
end if
return U(x)
```

Algorithm 4 Subroutine for Determining Minimal Assets

$path$ = "sequence of nodes on the feasible s,t-path"
L = "number of nodes in $path$
IA = "the vector of initial assets used in each mode $1 : m$"
EA = "the vector of excess assets used in each mode $1 : m$"
RA = "the vector of reduced assets used in each mode $1 : m$"
EM = "the vector of excess MTM used in each mode $1 : m$"
AM = "the vector of available MTM used in each mode $1 : m$"
IM = "the vector of initial MTM used in each mode $1 : m$"
RM = "the vector of reduced MTM used in each mode $1 : m$"
U = "the vector of UTE rates for each mode $1 : m$"
PA = "the vector of single-trip payloads (in tons) for each mode $1 : m$"
PR = "the vector of productivity rates for each mode $1 : m$"
BL = "the vector of block speeds for each mode $1 : m$"
AS = "the vector of max assets available for each mode $1 : m$"
for $j = 1 : m$ **do**
  for $i = 1 : L - 1$ **do**
    if mode of $path(i) \equiv$ mode of $path(i+1) \equiv j$ **then**
      $IM(j) \leftarrow IM(j) +$ arclength $(path(i), path(i+1))$
    end if
  end for
end for
for $i = 1 : m$ **do**
  $MTMasset(i) \leftarrow (U(i) * PA(i) * BL(i) * PR(i))/1,000,000$
  $AM(i) \leftarrow MTMasset(i) * AS(i) * Time$
end for
$EM \leftarrow AM - IM$
for $i = 1 : m$ **do**
  $EA(i) = floor(\frac{EM(i)}{AM(i)*Time})$
end for
for $i = 1 : m$ **do**
  $RA(i) = IA(i) - EA(i)$
end for
return RA

Algorithm 5 Subroutine for Generating Near-Efficient Solutions

```
if "lexicographic ordering is provided" then
    lexOrd = 1
else if "lexicographic ordering is not provided" then
    lexOrd = 0
end if
for i = 1 : IntNum + 1 do
    R = minT + i * IntSize
    x = f(R, MinAssets)
    if x = −1 then
        minDays = StValInc
    end if
    runDays = minDays + intSize * (i − 1)
    AssetMixVect = AssetMix(lexOrd, runDays)
    minAssetMixVect = FeasAssetMin(AssetMixVect)
end for
```

on a one thousand square mile grid. All parameter values are summarized in the Table 5.3.1.1:

5.3.1.1 An Example Problem. The results of the analysis are provided in tables 5.3 and 5.4 below. Notice that the minimum number of days required to provide the required capacity with the given assets is 127. This value is calculated using all available assets. Another feature of the methodology illustrated by the results is that the asset mix for the minimum number of days is less than all available assets. In fact, using bisection, this asset mix is minimal. In other words, a reduction of a single unit of any asset in the given asset mix results in an infeasible problem for a given value of T.

Two other concepts bear explanation. First, notice that costs of solutions 3, 4, and 5 in the asset/time example are identical. The cost expressed is for the number of MTM *used* rather than the number *available* for the given asset mix. The increase in time between solution one and solution two results in an increase in MTM provided on a per asset basis. Therefore, although the number assets in the

interval length (in days)	30
number of intervals	5
grid size (in square miles)	800
POE (source node)	1
POD (sink node)	50
number of millions of tons to move	5
arc density	90%
max assets available	[10,15,10]
cost per MTM (in $)	[100,110,120]
lexicographic ordering (3 = most important)	[3,2,1]
UTE rate (hours/day)	[15,15,15]
productivity (%)	[70,70,70]
block speed (mph)	[515,366,579]
payload (tons)	[85,22,135]
transfer costs (in $ per MT)	0.000001

Table 5.2: Small Experiment Parameters

asset mix has decreased, the same number of MTM are being utilized to provide the required capacity.

Secondly, notice that the cost increases in subsequent solutions in the assets/-time example while they are reduced in the cost/time example. This is explained by the lexicographic order when using the lexicographic order rather than cost as an objective component, the most expensive mode of transportation is given an objective component ordering as the *least* important or critical transportation mode. Therefore, the methodology seeks to minimize assets in the other two modes of transportation first which results in a greater amount of capacity being provided by the most expensive (but least important) transportation mode.

Using the generalized range variance developed in section 5.2, the range variance for these solutions is calculated as 0.0601 indicating that the near-efficient points are uniformly spaced in the objective space.

Now if we solve the same problem based upon minimizing cost rather than based upon a lexicographic ordering of the transportation modes we obtain the solutions in table 5.4.

solution	days	assets	\$ millions	path (original)	path (with modes)
1	127	[9,0,10]	177.99	[1,10,32,50]	$\overset{mode1}{[1,10]}\overset{mode3}{[10,32]}\overset{mode1}{[32,50]}$
2	157	[7,0,8]	178.09	[1,32,50]	$\overset{mode3}{[1,32]}\overset{mode1}{[32,50]}$
3	187	[5,0,10]	187.95	[1,10,50]	$\overset{mode1}{[1,10]}\overset{mode3}{[10,50]}$
4	217	[5,0,9]	187.95	[1,10,50]	$\overset{mode1}{[1,10]}\overset{mode3}{[10,50]}$
5	247	[4,0,8]	187.95	[1,10,50]	$\overset{mode1}{[1,10]}\overset{mode3}{[10,50]}$
6	277	[0,0,7]	188.03	[1,50]	$\overset{mode3}{[1,50]}$

Table 5.3: Multiobjective Lexicographic Order & Time Solution Results

solution	days	assets	\$ millions	path (original)	path (with modes)
1	127	[9,0,10]	177.99	[1,10,32,50]	$\overset{mode1}{[1,10]}\overset{mode3}{[10,32,50]}$
2	157	[10,0,7]	174.53	[1,44,50]	$\overset{mode1}{[1,44]}\overset{mode3}{[44,50]}$
3	187	[9,0,4]	170.76	[1,44,50]	$\overset{mode1}{[1,44]}\overset{mode3}{[44,50]}$
4	217	[9,0,2]	166.69	[1,44,50]	$\overset{mode1}{[1,32]}\overset{mode3}{[32,50]}$
5	247	[9,0,2]	164.60	[1,44,32,50]	$\overset{mode1}{[1,44]}\overset{mode3}{[44,32]}\overset{mode1}{[32,50]}$
6	277	[9,0,2]	163.22	[1,44,32,50]	$\overset{mode1}{[1,44]}\overset{mode3}{[44,32]}\overset{mode1}{[32,50]}$

Table 5.4: Multiobjective Cost & Time Solution Results

The range variance of this solution is calculated as 1613.09. This range variance indicates less uniformity in the spacing of the efficient solutions. However, the magnitude of the cost objective component is quite large numbering in the hundreds of thousands whereas the magnitude of the largest objective component of time in the cost/time example was only in the hundreds.

Each node in the network is given a unique identifier which indicates the corresponding node and mode number. Each node $v = i + n(k-1)$ where i is the node identifier in the original network and k is the number of modes of transportation in the problem [56]. For example, in a multimodal network with $n = 10$ nodes and $k = 3$ modes of transportation, node five in mode three would receive the unique identifier $v = 5 + 10(3-1) = 25$. With this node labeling method we describe one of the multimodal paths in the solutions above. The remaining paths can easily be obtained by the reader. The path in solution 1 in Table 5.3 consists of three shipping arcs and two mode transfer arcs. Arc $(1, 10)$ is in mode 1. At node 10 capacity is transferred to mode 3. Capacity is added to arc $(10, 32)$ in mode 3. At node 32 capacity is transfered to mode 1 and capacity is added to arc $(32, 50)$ completing the path.

5.4 Conclusions and Future Work

5.4.1 Conclusions. The methodology developed in this research provides insight into the interaction of time, transportation asset quantity requirements, and total cost. Since the RCSP is formulated based on a desire to generate a rough-cut capacity estimate, the data inputs required to complete this analysis are limited to data that are almost always available and will not need to be collected such as asset capacity, speed, utilization rate, and productivity among others.

Very little input is required on the part of the decision maker in order to complete the analysis. Once the distribution network is captured, and the engineering data for all transportation assets is obtained, analysis only requires three pieces of information from a decision maker: maximum number of transportation assets available in each mode, length of an interval (in days) between solutions, and number of intervals for which to generate solutions. The latter two items determine both the fidelity and solution space coverage.

This methodology provides an explicit analysis of the tradeoffs involved between cost, time, and transportation assets and provides a decision maker with the capability to determine a point on this continuum at which an acceptable balance of all three goals is achieved.

VI. Summary & Conclusions

This dissertation has made several original contributions to the field of Operations Research. First, a new methodology for modeling the multimodal RCCP using the RCSP was developed in Chapter 3 along with results showing the validity of the method using a binary integer programming formulation and a commercial solver.

Chapter 4 developed a new ACS metaheuristic for solving the RCSP which is a NP-Complete optimization problem. Results of this section show that the meta-heuristic is able to quickly find high quality feasible solutions to large instances of the RCSP.

Chapter 5 builds upon the modeling approach of Chapter 3 and the ACS metaheuristic of chapter 4 by developing a new metaheuristic for generating near Pareto solutions to the multimodal RCCP. The methodology considers the situation where cost and time are the objective components of interest and also considers a lexicographic objective ordering by which modes of transportation can be rank ordered according to importance or criticality. An extension of the range variance equation is developed to analyze the uniformity of near efficient solution spacing on the near Pareto front generated.

Future research should focus on generating larger instances of the RCSP for formulation as a BIP and solution by commercial solver software. The focus of this dissertation is on rough-cut capacity planning which seeks to answer the question: "Is there sufficient capacity in the system to meet transportation demands." Another area of future research is aggregate transportation planning. In production and inventory management this planning concept is used to answer the question: "How much should I produce in each planning period in order to meet demands across all planning periods. Aggregate transportation planning would ask an analogous question: "What quantities of freight should be transported in each period to meet demands across all planning periods." Aggregate and rough-cut capacity planning are closely related since the system capacity presents a constraint that must be considered in aggregate planning. Aggregate planning in transportation should consider crew scheduling, labor, warehousing, inventory, production, shipping, seasonal demand, etc.

The author believes this area of research to be particularly rich it topic areas, and timely considering the current need to reduce expenditures while maintaining levels of transportation capacity. One way to achieve this is through the application and extension of current Operations Research methods in order to improve efficiency and effectiveness of transportation planning systems.

VII. Appendix 1: FCD Experiment Results For ACS Testing

alpha	beta	max_iterations	num_ants	num_neighbors	q_0	rho	xi	percent over optimal	run time
-1	1	-1	1	-1	-1	-1	1	0.820803265	13.89314382
-1	1	1	-1	1	-1	1	1	0.004476521	11.47112901
1	-1	-1	1	1	-1	-1	-1	8.356031737	10.39651338
-1	-1	1	-1	1	-1	-1	1	0.004476511	13.6202956
-1	-1	-1	-1	-1	-1	-1	1	0.220307227	1.160888884
1	1	-1	-1	-1	1	-1	-1	0.004476511	1.182832176
1	1	-1	1	-1	-1	-1	-1	3.323575033	8.186534322
-1	-1	1	-1	-1	1	-1	1	0	12.6164134
-1	-1	1	1	-1	1	1	1	0	284.0650584
-1	1	-1	-1	1	-1	1	-1	0.004476511	1.924434949
0	0	0	-1	0	0	0	0	0.004476506	3.033284957
0	0	0	0	0	0	0	0	0.004476511	7.532166937
-1	1	1	1	1	1	-1	-1	0	351.2718527
-1	1	1	-1	1	1	1	-1	0	13.5170625
1	-1	-1	1	-1	1	1	1	0	12.07036425
0	0	1	0	0	0	0	0	4.07792742	59.80128759
-1	-1	1	1	1	-1	1	1	12.75990173	105.9970411
1	1	-1	1	1	-1	-1	1	0.004476516	11.03437797
1	-1	-1	1	-1	-1	1	-1	4.077927415	7.05146101
-1	-1	-1	1	-1	-1	-1	-1	4.07792743	8.662932021
1	-1	1	-1	1	-1	1	1	11.92272051	12.79150227
1	-1	-1	-1	1	1	1	1	0.004476511	1.158262433
1	-1	1	-1	1	1	-1	1	0	12.25825894
1	1	1	1	-1	-1	-1	1	0.82080326	296.1937443
-1	-1	-1	1	1	-1	-1	1	0.831086705	11.71672769
1	-1	-1	-1	-1	-1	1	1	4.07792742	1.040556272
1	1	1	-1	-1	1	-1	1	0	12.9834465
1	-1	-1	-1	1	1	-1	-1	0.004476511	1.172273152
-1	1	-1	1	-1	1	-1	-1	0	13.38079281
-1	-1	1	1	-1	-1	-1	1	3.323575038	364.9890439
1	0	0	0	0	0	0	0	3.338162012	9.002755658
1	1	-1	1	1	1	-1	-1	0	13.71408643
1	1	-1	-1	1	-1	-1	-1	18.49638851	1.772449152
-1	1	-1	1	-1	-1	1	-1	0.004476506	12.65459619
1	1	1	1	-1	-1	1	-1	0.004476511	287.9672545
1	1	-1	1	1	-1	1	-1	29.95927477	10.06518692
1	1	1	-1	1	1	-1	-1	0	13.88551646
1	-1	1	1	1	1	-1	-1	0	349.5637215
-1	-1	1	1	-1	-1	1	-1	3.323575033	390.7650016
-1	1	1	1	-1	-1	1	1	3.323575038	389.5671404
0	0	0	0	1	0	0	0	4.07792742	16.55031827
-1	1	1	1	-1	1	-1	1	0	279.400979
-1	1	1	1	1	-1	-1	1	20.14362899	113.3697083
-1	1	-1	1	1	-1	1	1	9.696973944	10.61463901
-1	1	-1	-1	-1	1	1	-1	0.004476511	1.092379199
-1	1	-1	-1	1	-1	-1	1	29.95927476	2.044553847
-1	1	1	-1	-1	-1	1	-1	26.19590945	11.50338954
1	1	1	-1	-1	-1	-1	-1	0.82080326	13.74440145
1	-1	-1	1	1	1	1	-1	0	13.28733065
1	1	1	1	1	-1	1	1	29.95927476	110.4218117
1	1	1	1	-1	1	1	1	0	286.3668183

Table 7.1: FCD Experimental Results Part 1

alpha	beta	max_iterations	num_ants	num_neighbors	q_0	rho	xi	percent over optimal	run time
1	1	-1	1	-1	-1	1	1	0.004476511	8.776136073
-1	1	-1	-1	1	1	-1	-1	0.004476511	1.188426881
1	-1	1	-1	1	1	1	-1	0	14.14519527
1	-1	-1	-1	-1	1	-1	1	0.004476511	1.168844644
0	0	0	0	0	0	1	0	4.07792743	8.870251945
1	-1	1	1	-1	1	1	-1	0	347.4683644
-1	-1	1	-1	-1	-1	-1	-1	4.725338158	7.689125853
1	-1	1	-1	-1	1	-1	-1	0	13.23266337
-1	-1	-1	-1	1	1	-1	1	0.004476511	1.118806301
-1	1	1	1	1	-1	1	-1	3.323575038	101.132139
1	1	-1	-1	-1	-1	1	-1	4.077927425	1.021030606
1	-1	1	1	1	-1	-1	1	8.356031737	90.26189041
0	0	0	0	0	-1	0	0	28.9359434	9.238635357
1	-1	-1	1	1	-1	1	1	8.356031742	10.70448409
1	-1	-1	1	1	1	-1	1	0	12.19686343
1	-1	1	-1	-1	-1	-1	1	0.82080326	10.67592725
1	-1	1	-1	-1	-1	1	-1	0.220307227	11.95604637
-1	1	-1	1	1	1	-1	1	0	12.27324157
1	-1	-1	-1	1	-1	-1	1	0.82080326	1.648543879
-1	-1	1	-1	-1	-1	1	1	1.63713001	10.19278707
-1	1	1	1	-1	1	1	-1	0	347.2821683
1	1	1	1	-1	1	-1	-1	0	348.9279031
-1	-1	-1	-1	1	-1	1	1	12.85493611	2.029011928
-1	-1	-1	1	-1	1	1	-1	0	13.89327317
0	0	-1	0	0	0	0	0	0.004476521	4.576830156
-1	-1	-1	1	-1	-1	1	1	4.077927425	6.010761982
-1	-1	1	1	1	1	1	-1	0	355.2176164
1	-1	1	1	1	1	1	1	0	283.1570151
0	0	0	0	0	0	0	1	0.004476506	8.650500413
1	-1	-1	-1	-1	-1	-1	-1	0.820803265	1.145741568
0	0	0	0	0	0	0	0	0.004476511	8.003210883
1	1	-1	-1	-1	-1	-1	1	1.637130005	1.113309233
-1	1	1	-1	1	-1	-1	-1	1.63713001	9.883301509
-1	1	1	-1	-1	1	1	1	0	12.58695421
1	1	-1	-1	-1	1	1	1	0.004476511	1.128043089
0	0	0	0	0	0	-1	0	0.004476516	18.16765092
1	1	-1	-1	1	1	1	-1	0.004476511	1.165215551
1	1	-1	-1	1	1	-1	1	0.004476511	1.147504641
1	1	-1	1	-1	1	-1	1	0	12.822396
0	-1	0	0	0	0	0	0	0.004476506	7.591606755
-1	-1	-1	-1	-1	-1	1	-1	0.82080326	1.111844522
1	-1	1	1	-1	-1	1	1	0.312667506	278.8550166
-1	-1	1	-1	-1	1	1	-1	0	13.66741485
-1	1	1	-1	-1	1	-1	-1	0	13.34852041
-1	1	1	-1	-1	-1	-1	1	1.637130005	8.078617813
1	-1	1	1	-1	1	-1	1	0	286.028231
-1	1	-1	-1	-1	-1	-1	-1	0.820803255	1.144141783
1	-1	-1	1	-1	1	-1	-1	0	13.61068865
-1	0	0	0	0	0	0	0	20.50392892	9.560084731
1	1	1	1	1	1	-1	1	0	267.8732819

Table 7.2: FCD Experimental Results Part 2

alpha	beta	max_iterations	num_ants	num_neighbors	q_0	rho	xi	percent over optimal	run time
-1	1	-1	1	-1	1	1	1	0	11.55188327
1	-1	1	1	1	-1	1	-1	0.004476511	100.8008896
-1	-1	-1	1	-1	1	-1	1	0	11.57078539
-1	-1	-1	1	1	1	-1	-1	0	12.44630337
1	-1	1	-1	1	-1	-1	-1	3.323575043	10.61495162
1	1	1	1	1	-1	-1	-1	8.356031737	86.57042836
-1	-1	1	1	1	-1	-1	-1	0.004476516	95.10207429
-1	-1	-1	1	1	1	1	1	0	11.72501128
-1	-1	1	-1	1	1	1	1	0	11.98009328
1	1	1	-1	1	1	1	1	0	12.22406884
-1	1	-1	-1	-1	-1	1	1	27.21905124	1.097110806
-1	-1	1	-1	1	-1	1	-1	36.7471159	12.69746448
-1	-1	1	-1	1	1	-1	-1	0	12.84176397
-1	1	1	-1	1	1	-1	1	0	11.85530815
1	-1	-1	-1	1	-1	1	-1	0.004476516	1.869312161
1	1	1	-1	-1	1	1	-1	0	13.04782186
1	1	1	-1	-1	-1	1	1	0.024175698	13.86908252
-1	-1	1	1	1	1	-1	1	0	263.5471229
-1	-1	-1	1	1	-1	1	-1	0.820803276	14.00132335
-1	1	1	1	1	1	1	1	0	262.6201033
-1	1	-1	1	1	-1	-1	-1	0.004476501	11.84573556
-1	-1	-1	-1	-1	1	1	1	0.004476511	1.141610316
-1	1	1	1	-1	-1	-1	-1	1.63713001	274.9328212
0	0	0	1	0	0	0	0	0.004476506	39.06677134
-1	1	-1	1	1	1	1	-1	0	12.48501304
1	-1	-1	1	-1	-1	-1	1	27.21905123	6.271068669
1	-1	-1	-1	-1	1	1	-1	0.004476511	1.139471357
-1	1	-1	-1	1	1	1	1	0.004476511	1.10176461
0	0	0	0	-1	0	0	0	4.959628171	10.77465139
1	1	1	-1	1	-1	1	-1	13.28078956	9.88961111
1	1	1	-1	1	-1	-1	1	14.50413045	11.16779772
1	1	-1	1	-1	1	1	-1	0	12.52054144
-1	1	-1	-1	-1	1	-1	1	0.004476511	1.104824077
1	1	-1	1	1	1	1	1	0	11.57956528
1	-1	1	1	-1	-1	-1	-1	4.077927425	468.4784276
1	1	-1	-1	1	-1	1	1	0.004476511	1.96316543
-1	-1	1	1	-1	1	-1	-1	0	323.6814468
-1	-1	-1	-1	-1	1	-1	-1	0.004476511	1.102508
1	-1	1	-1	-1	1	1	1	0	11.88593299
1	1	1	1	1	1	1	-1	0	317.8588666
-1	-1	-1	-1	1	-1	-1	-1	0.024175693	1.711611252
-1	-1	-1	-1	1	1	1	-1	0.004476511	1.142997225
0	0	0	0	0	1	0	0	0	16.77154163
0	0	0	0	0	0	0	-1	4.07792743	8.818615355
0	1	0	0	0	0	0	0	0.004476516	14.23597804

Table 7.3: FCD Experimental Results Part 3

Bibliography

1. "U.S. Census Bureau Economic Census 2007: Industry Snapshot", 2012.
2. Ahuja, R. K., T. L. Magnanti, and J. B. Orlin. *Network Flows*. Prentice Hall, Upper Saddle River, NJ, first edition, 1993.
3. Amiouny, S. V., J. J. Bartholdi, and J. H. Vande Vate. "Balanced Loading". *Operations Research*, 40(2):238–246, 1992.
4. Andersen, J., G. T. Crainic, and M. Christiansen. *Service Network Design with Management and Coordination of Multiple Fleets*. Technical Report CIRRELT-2007-21, 2007.
5. Aneja, Y., V. Aggarwal, and K. Nair. "Shortest Chain Subject to Side Conditions". *Networks*, 13:295–302, 1983.
6. Armacost, A. P. "Composite Variable Formulations for Express Shipment Service Network Design", 2000.
7. Armacost, A. P., C. Barnhart, and K. A. Ware. "Composite Variable Formulations for Express Shipment Service Network Design". *Transportation Science*, 36(1):1–20, 2002.
8. Asakura, Y. "Maximum capacity of road network constrained by user equilibrium conditions". *24th Annual Conference of the UTSG*. 1992. See page 4 of Yang et. al. for the bi=level programming model.
9. Asperen, E. V. "Essays on Port, Container, and Bulk Chemical Logistics Optimization", 2009.
10. Avella, P., M. Boccia, and A. Sforza. "A Penalty Function Heuristic for the Resource Constrained Shortest Path Problem". *European Journal of Operational Research*, 142(2):221–230, 2002.
11. Avella, P., M. Boccia, and A. Sforza. "Resource Constrained Shortest Path Problems in Path Planning for Fleet Management". *Journal of Mathematical Modeling and Algorithms*, 3:1–17, 2004.
12. B., Douglas West. *Introduction to Graph Theory*. Prentice Hall, Upper Saddle River, NJ, second edition edition, 2001.
13. Ballew, B. P. "The Distributors Three-Dimensional Pallet-Packing Problem: A Mathematical Formulation and Heuristic Solution Procedure", 2000.
14. Balseiro, S. R., I. Loiseau, and J. Ramonet. "An Ant Colony algorithm hybridized with insertion heuristics for the Time Dependent Vehicle Routing Problem with Time Windows". *Computers and Operations Research*, 38:954–966, 2011.

15. Bard, J. F. and J. T. Moore. "A Branch and Bound Algorithm for the Bilevel Programming". *SIAM Journal on Scientific and Statistical Computing*, 11(2):281–292, 1990.

16. Barnhart, C. and H. D. Ratliff. "Modeling Intermodal Routing". *Journal of Business Logistics*, 14(1):205–223, 1993.

17. Barton, R. *Simplified Ant Colony System Applied to the Quadrativ Assignment Problem*. Technical Report NRC/ERB-1123, Natioal Research Council Canada, 2005.

18. Beasley, J. and N. Christofides. "An algorithm for the resource constrained shortest path problem". *Networks*, 19:379–394, 1989.

19. Bitran, Gabriel R., Elizabeth A. Haas, and Arnoldo C. Hax. "Hierarchical Production Planning: A Single Stage System". *Operations research*, 29(4):717–743, July 1 1981.

20. Bodin, L. and B. Golden. "Classification in Vehicle Routing and Scheduling". *Networks*, 11:97–108, 1981.

21. Boland, N., J. Dethridge, and I. Dumitrescu. "Accelerated Label Setting Algorithms for the Elementary Resource Constrained Shortest Path Problem". *Operations Research Letters*, 34:58–68, 2006.

22. Bonabeau, E., M. Dorigo, and G. Theraulaz. *Swarm intelligence: from natural to artificial systems*. Oxford University Press, 1999.

23. Brewer, A. M., K. J. Button, and D. A. Hensher. *Handbook of Logistics and Supply Chain Management*. Elsevier, 1st edition, 2001.

24. Burden, R. L. and J. D. Faires. *Numerical Analysis*. Thompson Brooks/Cole, 8th edition, 2005.

25. Calhoun, K. M. "A Tabu Search for Scheduling and Rescheduling of Combat Aircraft", 2000.

26. Chabrier, A. *Vehicle routing problem with elementary shortest path based column generation*. Technical report, ILOG.

27. Coello, C. A. Coello, G. T. Pulido, and M. S. Lechuga. "Handling Multiple Objectives with Particle Swarm Optimization". *IEEE Transactions on Evolutionary Computation*, 8(3):256–279, 2004.

28. Cox, J. F. and J. H. Blackstone Jr. *APICS Dictionary*. APICS, Falls Church, VA, 9 edition, 1998.

29. Crainic, G. T. and K. H. Kim. *Intermodal Transportation*. Technical report, 2005.

30. Deckro, R. F. and J. E. Hebert. “Goal Programming Approaches to Solving Linear Decision Rule Based Aggregate Production Planning Models”. *IIE Transactions*, 16(4):308–315, 1984.

31. Derringer, G. and R. Suich. “Simultaneous Optimization of Several Response Variables”. *Journal of Quality Technology*, 12:214–219, 1980.

32. Dongarra, Jack J. *Numerical linear algebra for high-performance computers; by Dongarra, J. J.; Philadelphia : Society for Industrial and Applied Mathematics, c1998.; Series: Software, environments, tools; Description: xviii, 342 p. : ill. ; 25 cm.; Notes: Includes bibliographical references (p. 301-328) and index.; ISBN: 0898714281 (pbk.); OCLC No.: ocm39890502.*

33. Dorigo, M. “Optimization, Learning, and Natural Algorithms”, 1992.

34. Dorigo, M., M. Birattari, and T. Stutzle. “Ant Colony Optimization: Artificial Ants as A Computational Intelligence Technique”. *IEEE Computational Intelligence Magazine*, (November):28–39, 2006.

35. Dorigo, M. and L. M. Gambardella. “Ant Colonies for the Traveling Salesman Problem”. *BioSystems*, 43:73–81, 1997.

36. Dorigo, M., V. Maniezzo, and A. Colorni. “The Ant System: Optimization by a Colony of Cooperating Agents”. *IEEE Transactions on Systems*, 26(1):1–13, 1996.

37. Dorigo, M. and T. Stutzle. *Ant Colony Optimization*. The MIT Press, Cambridge, Massachusetts, 1st edition, 2004.

38. Dror, M. “Note on the complexity of the shortest path models for column generation in VRPTW”. *Operations Research*, 42(5):977–978, 1994.

39. Dumitrescu, I. and N. Boland. “Improved Preprocessing, Labeling and Scaling Algorithms for the Weight-Constrained Shortest Path Problem”. *Networks*, 42(3):135–153, 2003.

40. Ehrgott, M. *Multicriteria Optimization*. 2010.

41. Elimam, A. and D. Kohler. “Two Engineering Applications of a Constrained Shortest Path Model”. *European Journal of Operational Research*, 103:426–438, 1997.

42. Eppstein, D. “Finding the k shortest paths”. *SIAM Journal on Computing*, 28(2):652–673, 1999.

43. Feillet, D., P. Dejax, M. Gendreau, and C. Gueguen. “An Exact Algorithm for the Shortest Path Problem with Resource Constraints: Application to some Vehicle Routing Problems”. *Networks*, 142:216–229, 2004.

44. Gambardella, L. M. and M. Dorigo. "Ant-Q: A Reinforcement Learning Approach to the Traveling Salesman Problem". *Proceedings of the Twelfth International Conference on Machine Learning (ML-95)*, 252–260. 1995.

45. Gambardella, L. M. and M. Dorigo. "Solving Symetric and Asymetric TSP's by Ant Colonies". *IEEE International Conference on Evolutionary Computation*, 622–627. 1996.

46. Gan, R., Q. Guo, H. Chang, and Y. Yi. "Improved ant colony optimization algorithm for the traveling salesman problems". *Journal of Systems Engineering and Electronics*, 21(2):329–333, 2010.

47. Garcia, R. "Resource Constrained Shortest Paths and Extensions", 2009.

48. Garey, R. G. and D. S. Johnson. *Computers and Intractability: A guide to the Theory of NP-Completeness*. W.H. Freeman and Company, 1979.

49. Ge, Y. E., H. M. Zhang, and W. H. K. Lam. "Network Reserve Capacity Under the Influence of Traveler Information". *Journal of Transportation Engineering*, 129(3):262–270, 2003.

50. Goss, S., S. Aron, J. L. Deneubourg, and J. M. Pasteels. "Self-Organized Shortecust in the Argentine Ant". *Naturwissenschaften*, 76:579–581, 1989.

51. Gueguen, C., P. Dejax, M. Dror, and M. Gendreau. *An exact algorithm for the elementary shortest path problem with resource constraints*. Technical Report G98-04-A, Laboratoire Productique Logistique, 1998.

52. Guinipero, L. C., T. E. Yoon, and S. Brudvig. "A Decade of SCM Literature: Past, Present, and Future Implications". *Journal of Supply Chain Management*, 2008.

53. Haghani, A. and O. Sei-Chang. "Formulation and Solution of a Multi-Commodity, Multi-Modal Network Flow Model for Disaster Relief Operations". *Transportation Research Part A*, 30(3):231–250, 1996.

54. Handler, G. and I. Zang. "A Dual Algorithm for the Constrained Shortest Path Problem". *Networks*, 10:293–310, 1980.

55. Hartlage, R. B. and J. D. Weir. "An Ant Colony System for Solving the Resource Constrained Shortest Path Problem (Working Paper)", 2011.

56. Hartlage, R. B. and J. D. Weir. "Rough-Cut Capacity Planning in Multimodal Freight Transportation Networks (Working Paper)", 2011.

57. Hassin, R. "Approximation Schemes for the Restricted Shortest Path Problem". *Mathematics of Operations Research*, 17(1):36–42, 1992.

58. Hax, A. C. and D. Candea. *Production and Inventory Management*. 1983. ISBN 0-13-724880-6.

59. Hoffman, W. and R. Pavley. "A Method for the Solution of the Nth Best Path Problem". *Journal of the Association for Computing Machinery*, 6(4):506–514, 1959.

60. Holt, C. C., F. Modigliani, and J. F. Muth. "Derivation of a Linear Decision Rule for Production and Employment". *Management Science*, 2(2):159–177, 1956.

61. Holt, C. C., F. Modigliani, and H. A. Simon. "A Linear Decision Rule for Production and Employment Scheduling". *Management Science*, 2(1-30), 1955.

62. Hove, J. C. Van. "An Integer Program Decomposition Approach to Combat Planning", 1998.

63. Hu, Y., Y. Liu, and W. Liu. "Multi-Constrained Shortest Path Model and Solution with Improved Ant Colony Algorithm". *Traffic and Transportation Studies*, 501–510, 2010.

64. Irnich, S. and G. Desaulniers. "Shortest Path Problems with Resource Constraints". *Les Cahiers du GERAD*.

65. Irnich, S. and D. Villeneuve. *The Shortest Path Problem with Resource Constraints and k-cycle Elimination for k ¿= 3*. Technical Report G-2003-55, Les Cahiers du GERAD, 2003.

66. Jacobs, F. R., W. Berry, D. Whybark, and T. E. Vollmann. *Manufacturing Planning and Control for Supply Chain Management*. McGraw-Hill Higher Education, 6th edition, 2011. ISBN ISBN-10 0-07-337782-1, ISBN-13 978-0-07-337782-7.

67. Jepsen, M., B. Petersen, and S. Spoorendonk. *A branch-and-cut Algorithm for the Elementery Shortest Path Problem with Resource Constraints*. Technical report, DIKU University of Copenhagen, 2005.

68. Ji, Junzhong, Renbing Hu, Hongxun Zhang, and Chunnian Liu. "A hybrid method for learning Bayesian networks based on ant colony optimization". *Applied Soft Computing*, 11(4):3373–3384, 6 2011.

69. Jr., E. E. Adam and R. J. Ebert. *Production and Operations Management*. Prentice Hall, Englewood Cliffs, NJ, 5th edition, 1992.

70. Kaluzny, B. L. and R. H. A. D. Shaw. *Optimal Aircraft Load Balancing*. Technical Report DRDC CORA TR 2008-004, Defense Research and Development Canada, Centre for Operational Research and Analysis, 2008.

71. Kasikitwiwat, P. and A. Chen. "Analysis of Transportation Network Capacity Related to Different System Capacity Concepts". *Journal of the Eastern Asia Society for Transportation Studies*, 6:1439–1454, 2005.

72. Katoh, N., T. Ibaraki, and H. Mine. “An efficient algorithm for k shortest simple paths”. *Networks*, 12:411–427, 1982.

73. Kennedy, K. “An Analysis of Multiple Layered Networks”, 2003.

74. Kennedy, K. “Synthesis, Interdiction, and Protection of Layered Networks”, 2009.

75. Kjerrstrom, J. “The Resource Constrained Shortest Path Problem”. 2001.

76. Konings, J. W. “Integrated centres for the transshipment, storage, collection, and distribution of goods”. *Transport Policy*, 3(1):3–11.

77. Lasdon, L. S. *Optimization Theory for Large Systems*. Macmillan, 1st edition, 1970.

78. Lorenz, D. H. and D. Raz. “A Simple Efficient Approximation Scheme for the Restricted Shortest Path Problem”. *Operations Research Letters*, 28(5):213–219, 2001.

79. Macharis, C. and Y. M. Bontekoning. “Opportunities for OR in Intermodal Freight Transport Research: A Review”. *European Journal of Operational Research*, 153:400–416, 2004.

80. Martins, G. H. A. “Packing in Two and Three Dimensions”, 2003.

81. McNabb, D. J. “US Transportation Command 2009 Strategic Plan”, 2009.

82. Mehlhorn, K. and M. Ziegelmann. *Resource Constrained Shortest Paths*. Technical report.

83. Meixell, M. J. and M. Norbis. “A review of the Transportation Mode Choice and Carrier Selection Literature”. *The International Journal of Logistics*, 19(2):183–211, 2008.

84. Merkle, D., M. Middendorf, and H. Schmeck. “Ant Colony Optimization for Resource-Constrained Project Scheduling”. *IEEE Transactions on Evolutionary Computation*, 6(4):333–346, 2002.

85. Monaco, M. F., L. Moccia, and M. Sammarra. “Operations Research for the Management of a Transhipment Container Terminal: The Gioia Tauro Case”. *Maritime Economics and Logistics*, (11):7–35, 2009.

86. Moore, J. T. and J. F. Bard. “The Mixed Integer Linear Bilevel Programming Problem”. *Operations Research*, 38(5):911–921, 1990.

87. Morlok, E. K. and D. J. Chang. “Measuring Capacity Flexibility of a Transportation System”. *Transportation Research Part A*, 38:405–420, 2004.

88. Morlok, E. K. and S. P. Riddle. “Estimating the Capacity of Freight Transportation Systems”. *Transportation Research Record*, (1653):1–8, 1999.

89. Mullen, R. J., D. Monekosso, S. Barman, and P. Remagnino. “A review of ant algorithms”. *Expert Systems with Applications*, 36:9608–9617, 2009.

90. Multiple. “USTRANSCOM Brief History (Official)”.

91. Multiple. *Air Mobility Planning Factors (AFPAM 10-1403)*. Technical report, US Air Force, 2003.

92. Myers, R. H., D. C. Montgomery, and C. M. Anderson-Cook. *Response Surface Methodology: Process and Product Optimization using Designed Experiments*. John Wiley and Sons, Inc., 3rd edition, 2009.

93. jin Nam, Sang and Rasaratnam Logendran. “Aggregate Production Planning - A Survey of Models and Methodologies”. *European Journal of Operational Research*, 61:255–272, 1992.

94. Nasiri, M. M. “A Pseudo Particle Swarm Optimization for the RCPSP”. *The International Journal of Advanced Manufacturing Technology*, 2012.

95. Nielsen, C. A., A. P. Armacost, C. Barnhart, and S. E. Kolitz. “Network Deign Formulations for Scheduling U.S. Air Force Channel Route Missions”. 39:925–943, 2004.

96. Park, M. and A. Regan. “Capacity Modeling in Transportation: A Multimodal Perspective”. *Transportation Research Record: A Multimodal Perspective*, 1906:97–104, 2005.

97. Pinedo, M. L. *Scheduling: Theory Algorithms and Systems*. 2008.

98. Pittman, S. D., B. B. Bare, and D. G. Briggs. “Hierarchical Production Planning in Forestry Using Price-Directed Decomposition”. *Canadian Journal of Forestry Research*, 37:2010–2021, 2007.

99. Ren, Zhi-Gang, Zu-Ren Feng, Liang-Jun Ke, and Zhao-Jun Zhang. “New ideas for applying ant colony optimization to the set covering problem”. *Computers and Industrial Engineering*, 58(4):774–784, 5 2010.

100. Ribeiro, C. and M. Minoux. “A Heuristic Approach to Hard Constrained Shortest Path Problems”. *Discrete Applied Mathematics*, 10:125–137, 1985.

101. Rousseau, L. M, F. Focacci, M. Gendreau, and G. Pesant. *Solving VRPTWs with constraint programming based column generation*. Technical Report CRT-2003-10, Center for Research on Transportation, 2003.

102. Russ, B. F., T. Yamada, J. T. Castro, and H. Yasukawa. “Optimizing the Design of Multimodal Freight Transport Network in Indonesia”. *Journal of the Eastern Asia Society for Transportation Studies*, 6:2894 – 2907, 2005.

103. Schott, J. A. “Fault Tolerant Design using Single and Multicriteria Genetic Algorithm Optimization”, 1995.

104. Sun, Y., M. A. Turnquist, and L. K. Nozick. "Estimating Freight Transportation System Capacity, Flexibility, and Degraded-Condition Performance". *Transportation Research Record*, 1966:80–87, 2006.

105. Systematics, Inc Cambridge. *NCHRP Report 399: Multimodal Corridor and Capacity Analysis Manual*. Technical report, Transportation Review Board, National Research Council, 1998.

106. Unnikrishnan, A. and S. T. Waller. "Freight Transportation Network Design Problem for Maximizing Throughput under Uncertainty". *Transportation Research Record*, 2090:105–114, 2009.

107. USTRANSCOM. *USTRANSCOM 2011 Strategic Plan*. Technical report, 2011.

108. Vis, I. F. A. and R. Koster. "Transhipment of Containers at a Container Terminal: An Overview". *European Journal of Operational Research*, (147):1–16, 2003.

109. Vollmann, T. E., W. L. Berry, and D. C. Whybark. *Manufacturing Planning and Control Systems*. McGraw Hill, 4th edition, 1997.

110. Wang, Z. and J. Crowcroft. "Quality-of-Sevice Routing for Supporting Multimedia Applications". *IEEE Journal on Selectede Areas in Communications*, 14(7):1228–1234, 1996.

111. White, T. *An Expert Assessment of Stigmergy*. Technical Report DRDC CR 2005-004, Defense Research and Development Canada, 2005.

112. Woxenius, J. *Intermodal Transshipment Technologies: An Overview*. Technical report, 1998.

113. Xue, G. "Primal-Dual Algorithms for Computing Weight-Constrained Shortest Path and Weight-Constrained Minimum Spanning Trees". *19th IEEE International Performance, Computing, and Communications Conference (IPCCC)*, 271–277. 2000.

114. Yagmahan, Betul and Mehmet Mutlu Yenisey. "A multi-objective ant colony system algorithm for flow shop scheduling problem". *Expert Systems with Applications*, 37(2):1361–1368, 3 2010.

115. Yang, H., M. G. H. Bell, and Q. Meng. "Modeling the Capacity and Level of Service of Urban Transportation Networks". *Transportation Research Part B*, 34:255–275, 2000.

116. Zhang, H., H. Li, and C. M. Tam. "Particle Swarm Optimization for resource-constrained project Scheduling". *International Journal of Project Management*, 24:83–92, 2006.

117. Zhu, X. and W. E. Wilhelm. "A three-stage approach for the resource-constrained shortest path as a sub-problem in column generation". *Computers and Operations Research*, 2011.

118. Zitzler, E., M. Laumanns, and L. Thiele. "SPEA2: Improving the Strength Pareto Evolutionary Algorithm". *Evolutionary Methods for Design Optimization and Control with Applications to Industrial Problems*. 2001.

119. Zografos, K. and A. Regan. "Current challenges for intermodal freight transport and logistics in Europe and the United States". *Transportation Research Record*, (1873):70–78, 2004. Doi: pmid:.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704–0188

1. REPORT DATE *(DD–MM–YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From — To)*
30–09–2012	Doctoral Dissertation	Aug 2009 — Sep 2012

4. TITLE AND SUBTITLE

Rough-Cut Capacity Planning in Multimodal Freight Transportation Networks

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Hartlage, Robert B., Maj, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT/DS/ENS/12-03

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

United States Transportation Command/Joint Distribution Process Analysis Canter (USTRANSCOM/TCAC)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approval for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

A main challenge in transporting cargo for United States Transportation Command (USTRANSCOM) is in mode selection or integration. Demand for cargo is time sensitive and must be fulfilled by an established due date. Since these due dates are often inflexible, commercial carriers are used at an enormous expense, in order to fill the gap in organic transportation asset capacity. This dissertation develops a new methodology for transportation capacity assignment to routes based on the Resource Constrained Shortest Path Problem (RCSP). Routes can be single or multimodal depending on the characteristics of the network, delivery timeline, modal capacities, and costs. The difficulty of the RCSP requires use of metaheuristics to produce solutions. An Ant Colony System to solve the RCSP is developed in this dissertation. Finally, a method for generating near Pareto optimal solutions with respect to the objectives of cost and time is developed.

15. SUBJECT TERMS

Network Optimization, Mathematical Modeling, Ant Colony System, Metaheuristics

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON / 19b. TELEPHONE NUMBER *(include area code)*
U	U	U	UU	127	Dr. Jeffery D. Weir / (937) 255–3636, ext 4523; jeffery.weir @ afit.edu

Standard Form 298 (Rev. 8–98)
Prescribed by ANSI Std. Z39.18